Volume 67, Part 1 *February 1976*

CUH- H06181- 39-P2230

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by A. D. B. CLARKE
with J. BROWN, R. CONRAD, N. F. DIXON,
G. W. GRANGER, I. MARTIN, R. L. REID,
B. TIZARD *and* P. B. WARR

39

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET, NEW YORK, N.Y. 10022

Assistant Editor: D. I. WILLIAMS

Book Review Editor: D. C. KENDRICK

Press Editor: MI

Vol. 67
pt. 1–4

Preparation of Pape [illegible] *sychology*

In submitting papers for publicatio [illegible] *Journal of Psychology* contributors [illegible] observe the following requirements.

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.
It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of articles published from Vol. 56 (1965). A brief summary, suitable for use as an abstract, should precede the article.

Manuscripts should be submitted in duplicate

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure [illegible] ould be numbered and on a separate sheet with its approximate position indicated in the typescript.
Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.
The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.
Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the author's names and corresponding to an alphabetical list at the end of the paper. This list must give author's names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–248.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Année psychol.* **50**, 83–96.

World list abbreviations are used for journals.
The Oxford Dictionary is taken as the authority on spelling.

The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

© The British Psychological Society 1976

Permissions

For permission to reproduce material from *The British Journal of Psychology*, please apply to the London or New York office of Cambridge University Press.

ISI Tear Service, 325 Chestnut Street, Philadelphia, Pennsylvania 19106, USA is authorized to supply single copies of separate articles for private use only.

Subscriptions

The British Journal of Psychology is published quarterly.

The subscription price (including postage) of volume 67, 1976, is £20.00 net (US $60.00 in the USA and Canada). Single parts costs £6.00 each (US $18.00 in the USA and Canada).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by A. D. B. CLARKE
with J. BROWN, R. CONRAD, N. F. DIXON,
G. W. GRANGER, I. MARTIN, R. L. REID,
B. TIZARD *and* P. B. WARR

Assistant Editor: D. I. WILLIAMS
Book Reviews Editor: D. C. KENDRICK
Press Editor: MICHELE BENJAMIN

VOLUME 67, 1976

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE · LONDON · NEW YORK

PUBLISHED BY
THE SYNDICS OF THE UNIVERSITY PRESS
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
Bentley House, 200 Euston Road, London NW1 2DB
32 East 57th Street, New York, N.Y. 10022

© The British Psychological Society 1976

P2230

Printed in Great Britain at the University Printing House, Cambridge

CONTENTS OF VOLUME 67

PART 1 FEBRUARY 1976

PAGE

Quantity conceptions in university students: another look. By J. A. ROWELL and V. J. RENNER 1

The golden section hypothesis. By J. BENJAFIELD and J. ADAMS-WEBBER . 11

Inspection time and measured intelligence. By T. NETTELBECK and M. LALLY 17

Implicit-shape constancy as a factor in pictorial perception. By J. B. DERĘGOWSKI 23

Some evidence for 'speech' as an acoustic feature. By JOHN MORTON and SUSAN M. CHAMBERS 31

Extraversion and reinforcement in verbal operant conditioning. By B. S. GUPTA 47

Extraversion, activation and the recall of prose. By MICHAEL W. EYSENCK . 53

The recall of concrete and abstract sentences as a function of interpolated task. By GRAHAM DAVIES and JOHN PROCTOR 63

Memory deficits in early infantile autism: some similarities to the amnesic syndrome. By JILL BOUCHER and ELIZABETH K. WARRINGTON . . . 73

Constraints on classificatory skills in young children. By ROBIN CAMPBELL, MARGARET DONALDSON and BRIAN YOUNG 89

The influence of context on children's use of identifying expressions and references. By DAVID A. WARDEN 101

Book reviews 113

Other publications received 133

PART 2 MAY 1976

Acquired sensory control of satiation in man. By D. A. BOOTH, M. LEE and CHRISTINE MCALEAVEY 137

Syntactic construction as retrieval plans. By GERARD KEMPEN 149

Recency in free recall: STM plus LTM or STM alone? By P. E. MORRIS and J. RIDGWAY 161

Semantic constraints and sentence perception. By E. STEPHAN NEWSTEAD . 165

Effects of novelty and oddity on visual selective attention. By D. E. BERLYNE and JOYCE DITKOFSKY 175

PAGE

Conditioning afterimages: a procedure minimizing the extinction effect of normal test trials. By PETER DAVIES 181

The role of information-processing capacity in the conserving operation. By VERNON HAMILTON and GILLES LAUNAY 191

Reproduction of Kohs-type figures by Ghanaian children: orientation errors revisited. By GUSTAV JAHODA 203

Were Cattell's 'personality sphere' factors correctly identified in the first instance? By EDGAR HOWARTH 213

How representative are concept attainment experiments? By N. E. WETHERICK and R. L. DOMINOWSKI 231

Parent–offspring resemblances in intelligence: theories and evidence. By MICHAEL MCASKIE and ANN M. CLARKE 243

Book reviews 275

Other publications received 297

PART 3 AUGUST 1976

Thought and the brain. By O. L. ZANGWILL 301

The effects of stimulus attributes upon latency of word recognition. By JOHN T. E. RICHARDSON 315

Semantic interference from unattended printed words. By GEOFFREY UNDERWOOD 327

Semantic coding versus the stimulus suffix. By D. SALTER, G. SPRINGER and L. BOLTON 339

Concurrent conditioning of evaluative meaning and imagery. By JEFFREY M. LOHR 353

A failure to explain the effects of false heart-rate feedback on affect by induced changes in physiological response. By E. W. THORNTON and PATRICIA J. HAGAN 359

The electrodermal component of the orienting response in blind and deaf individuals. By D. CARROLL and P. G. SURTEES 367

Frequency, conceptual structure and pattern recognition. By J. G. WOLFF . 377

Qualitative analysis of the performance of introverts and extraverts on standard progressive matrices. By VIDHU MOHAN and DALIP KUMAR . 391

Cognitive strategies and multidimensional scaling. By C. O. FRASER . . 399

PAGE

Right- and left-handed counting of Braille dots in subjects unaccustomed to Braille. By D. H. MYERS 407

The behavioural effects of single and combined stressors: a test of arousal theory. By P. M. LIND 413

The influence of reversibility and truncation on the interpretation of the passive voice by young children. By MARGARET HARRIS . . . 419

Stimulus processing and recognition memory in children. By DONALD H. SYKES 429

Book reviews 439

Other publications received 457

PART 4 NOVEMBER 1976

A relationship between Eysenck's P scale and change in muscle action potentials with attention to perceptual tasks. By JOHN HINTON and BRIAN CRASKE 461

Extraversion, neuroticism and strength of the nervous system. By JEAN-YVES FRIGON 467

Oculomotor adaptation to prisms: complete transfer between eyes. By MARTIN CRAWSHAW and BRIAN CRASKE 475

Rationalization in a reasoning task. J. ST B. T. EVANS and P. C. WASON . 479

Procedures for investigating imagery and the distinction between primary and secondary memory. JOHN T. E. RICHARDSON 487

The relation of reported imagery to cognitive performance. A. J. DURNDELL and N. E. WETHERICK 501

Visual behaviour in social interaction: the validity of eye-contact assessments under different conditions of observation. By BERNARD RIMÉ and LEX McCUSKER 507

A comment on Meddis' 'A simple two-group test for matched scores with unequal cell frequencies'. By R. N. SYKES 515

A critical note on Quinton and Fellows' observation of reasoning strategies. J. ST B. T. EVANS 517

The role of introspection in problem-solving research: a reply to Evans. By BRIAN J. FELLOWS 519

Rehearsal span versus memory span: what does this comparison tell us about verbal memory? By LAWRENCE PORTER 521

PAGE
The effect of embedded words in a brief visual display. DAVID E. SOMEKH 529
The components of the Poggendorff illusion. By R. H. DAY and R. G. DICKINSON 537
The effect of task complexity and sequence on rule learning and problem solving. By J. SWELLER 553
Responses to homophones as a function of cue words on the unattended channel. By S. H. A. HENLEY 559
The relative difficulty position function. By VICTOR M. AGRUSO, JR and MARK D. RECKASE 569
The difficulty of an implication task. By DAVID S. BRÉE and GEORGE COPPENS 579
A coordination of hand preference and skill replicated. By MARIAN ANNETT 587

Book reviews 593
Other publications received 609
Editorial acknowledgement 615

Br. J. Psychol. (1976), 67, 1, *pp.* 1–10
Printed in Great Britain

QUANTITY CONCEPTIONS IN UNIVERSITY STUDENTS: ANOTHER LOOK

BY J. A. ROWELL AND V. J. RENNER
Department of Education,
University of Adelaide, South Australia

On hundred and thirty-eight post-graduate students enrolled in a Diploma in Education course were tested for their ability to conserve weight and volume. A methodological criticism of the earlier work of Elkind (1962) and Towler & Wheatley (1971) is offered which casts doubt on the extremely high proportion of non-conservers of volume reported in those studies. Nevertheless, using a modified version of Elkind's methodology, 18 per cent of our sample were non-conservers of volume, a result providing some support for the generality of Elkind's findings. The results are discussed under the four headings of horizontal *décalage*, sex differences, possible effects of pre-graduate courses of study and educational implications. The importance of the individual as constructor and interpreter of his own experience is emphasized.

In 1962 Elkind brought attention to the rather remarkable fact that only 58 per cent of a sample of 239 first-year university and college students had an abstract conception of volume. Since that time there appears to have been a single restricted attempt – limited to one sex only – to test the validity of these findings by a direct replication of the Elkind study. In 1971 Towler & Wheatley found that 61 per cent of a sample of female university students enrolled in a first-year course in mathematics did not conserve volume, i.e. did not maintain invariance of volume across transformation in shape. McKinnon & Renner (1971) reported 27 per cent of a sample of 131 first-year university students failed to conserve volume and/or quantity. However, the direct equivalence of their results is not known, as details of their methodology were not provided.

In 1973, as part of a larger study carried out with the cooperation of postgraduate Diploma in Education students at the University of Adelaide, South Australia, a second attempt was made to test Elkind's original findings. It was decided also to examine the differences, if any, between the sexes and between the effects of pre-graduate courses of study leading to the degrees of B.Sc. and B.A.

METHOD

One hundred and thirty-eight post-graduate students enrolled in the Diploma in Education course at the University of Adelaide, South Australia, took part in the study. Forty-six were science graduates, 32 men (median age 22, age range 20–29), and 14 women (median age 21, age range 20–25), and 92 were graduates in the arts, 25 men (median age 22, age range 20–28), and 67 women (median age 22, age range 20–42). All ages were to the nearest whole year at the time of taking the test. For this reason both medians and ranges have been given in whole numbers also. Eight science graduates (17·4 per cent of that group) and 17 arts graduates (18·5 per cent of that group) had an honours degree.

Procedure

The method of administration closely followed that outlined by Elkind (1961, 1962) and utilized by Towler & Wheatley (1971). A group method was used; in our case the groups were tutorial groups consisting of 8–10 students.

In an attempt to avoid verbal misunderstandings the two balls of plasticine used in the experiment were weighed on a beam balance in front of the students, after first checking to their satisfaction that the balance was operating correctly. Students were encouraged to make any alterations they wished to the plasticine balls until they appeared to be of equal weight within the limits of accuracy allowed by the balance. When this had been achieved and each member of the group had asserted the equality in weight of the two plasticine balls, they were assured that the experiments which followed would involve no trickery of any sort and that the two balls were, as far as we could ensure it, identical in every respect.

Questions were asked concerning weight and volume only. Unlike the Elkind (1962) and Towler & Wheatley (1971) studies, no questions were asked concerning quantity for two reasons. First, because of the strong probability that all students would be conservers of quantity (as found by Towler & Wheatley, 1971) and hence might become bored by the 'obviousness' of the proceedings. In this respect, although a subjective impression, it does seem worth recording that the facial expressions of many students could readily be interpreted as incredulity that the experimenters should ask such 'obvious' questions when weight was the subject of investigation. However, a number of these expressions altered considerably when the concept of volume was introduced. The second reason for omission of questions concerning quantity conservation was the time limitation of the tutorial sessions.

For both weight and volume questions were asked concerning identity, prediction and judgement, and explanations were asked for in each case. Following Elkind (1962), a test was considered passed only if a student made no unambiguous error in any section of the problem, including explanation. The questions, to which students wrote their answers on a sheet of paper, were as follows:

A. Weight

1. Do the balls of plasticine weigh the same? (Identity question.)
2. Explain your answer.
3. Suppose I made one of the balls into a sausage shape. Would the two pieces of plasticine, the ball and the sausage, weigh the same? (Prediction.)
4. Explain your answer.

At this point one of the balls of plasticine was rolled into a sausage shape (the length of the sausage was maintained at a constant 9–10 in. for each tutorial group).

5. Do the two pieces of plasticine, the ball and the sausage, weigh the same? (Judgement.)
6. Explain your answer.

B. Volume

Before asking any questions about volume the sausage-shaped piece of plasticine was rolled back into a ball as closely as possible equivalent to the other ball of plasticine. Questions were then asked as follows:

1. Do the balls of plasticine have the same volume? (Identity.)
2. Explain your answer.
3. Suppose I made one of the balls into a sausage shape. Would the two pieces of plasticine, the ball and the sausage, have the same volume? (Prediction.)
4. Explain your answer.

At this point one of the balls of plasticine was rolled into a sausage shape.

5. Do the two pieces of plasticine, the ball and the sausage have the same volume? (Judgement.)
6. Explain your answer.
7. Suppose I rolled the sausage-shaped piece of plasticine back into a ball would the two balls of plasticine have the same volume? (Prediction.)

8. Explain your answer.
9. What would happen to the water levels if the ball and the sausage were placed into identical large glasses filled equally high with water? Assume the plasticine is completely immersed in the water in each case and that no spillage of water occurs. (These last points, not included in the Elkind (1962) and Towler & Wheatley (1971) studies, were included to avoid any possible ambiguity in the situation.)
10. Explain your answer.

Questions 7 and 8 were not part of the Elkind (1962) or Towler & Wheatley (1971) studies but were included here as a check on the consistency of arguments advanced by individuals, particularly in relation to density. This seemed necessary, as Elkind (1961, 1962) judged as non-conservers some individuals giving answers which appear to the present authors to be strongly indicative of conservation; for example, Elkind (1962) quotes as a non-conservation answer: 'Allen (20). One ball can be made more compact than the other one; therefore they are not the same volume.' Presumably this is only a non-conservation answer if density is being confused in some way with shape. Whether this is the case can readily be decided by answers given to questions 7 and 8 included here. Although they are not as explicit, Towler & Wheatley (1971) also refer to many 'erroneous explanations' in terms of density arguments, and, as they offer no criticism of Elkind's work, presumably accept his arguments.

RESULTS

With one exception all students including those who gave non-conservation answers, stated that the two balls of plasticine weighed the same at the start of the experiment. The single non-agreement was based on the information already provided by the experimenters that weight equivalence of the two balls of plasticine was within the limits of accuracy of the balance and did not mean absolute equivalence. The negative response, given by a female arts graduate, stated 'No. The balance is only approximate'. However, having shown she recognized the distinction, the same student then continued her answers without further reference to it. For example, when asked to predict whether the two pieces of plasticine would weigh the same if one were rolled into a sausage she replied: 'Yes. The solid mass will weigh the same no matter what the shape.'

Only one student, a male science graduate, made statements which could be interpreted as denying conservation of weight. Having agreed that the two balls of plasticine weighed the same at the start of the experiment the replies to the prediction and judgement questions involving transformation of shape were: 'No, the one with the sausage should weight [*sic*] heavier.' And the explanation given was 'Because of the action of the arms on the centre of gravity in the elongated one.' It is *possible* to explain this without resorting to lack of conservation by saying that the student considered the sausage-shaped piece of plasticine now occupied a different position on the balance scale such that its centre of gravity was further from the fulcrum. This does seem to be bending over backwards to explain the statement however, and it seems more likely to the authors either that the student was confused over the position of the centre of gravity of the plasticine when rolled into a sausage, or that it was a genuine case of non-conservation. The former possibility is given more credence by the fact that, in our marking scheme, the student was a conserver of volume, utilizing a density argument to account for any possible volume change brought about by the act of deformation – 'The volume of the sausage will depend on the force by which it is being made into a sausage.' However a degree of doubt

Table 1. *Numbers and percentages of students giving conservation and non-conservation type answers to questions on volume*

	Science graduates		Arts graduates	
	Male	Female	Male	Female
Conservers	90·6 % $n = 29$	78·6 % $n = 11$	88 % $n = 22$	71·6 % $n = 48$
Non-Conservers	9·4 % $n = 3$	21·4 % $n = 3$	12 % $n = 3$	28·4 % $n = 19$
n	$n = 32$	$n = 14$	$n = 25$	$n = 67$

s attached to the conserver decision in this case also by another anomaly in his thought processes; having indicated the possible decrease in volume of the sausage due to compaction he subsequently inferred it had a *larger* volume than the ball in his comments on what would happen to the water levels in the two identical beakers.

An inspection of Table 1 allows two important comments to be made. First, for this sample, there is a substantial percentage of students (approximately 13 per cent) who graduate in science subjects and yet do not appreciate that the volume of a constant mass of plasticine is independent of its shape. Secondly, independent of the pregraduate courses of study followed (for the 2 × 2 contingency table arts/sciences–conservers/non-conservers, $\chi^2 = 1{\cdot}62$, a non-significant value) a significantly higher percentage of men than women have an abstract conception of volume ($\chi^2 = 4{\cdot}74$, $P < 0{\cdot}05$) – an observation in agreement with that reported by Elkind (1961, 1962).

With one exception, all non-conservers of volume indicated that they believed the volumes of the two balls of plasticine were the same initially, or would be if they were *precisely* the same shape, and that if the sausage were rolled back into a ball after the experiment they would again be of the same volume. The one exception, a female arts graduate, denied that solids had volume. Her answer to the identity question was, 'No – volume is a measure of space or capacity. Balls are solid entities.' And to the judgement question after rolling one ball into a sausage, 'No – a solid object does not have volume regardless of shape.' For her volume was a measure of 'empty' space, a property of receptacles – unless something could be poured into it the term volume was inapplicable. Interestingly, and significantly, no student using a compression argument to explain change in volume accompanying change in shape predicted equality of volume in the event of the sausage again being rolled into a ball (answer to question 7). This evidence appears to refute some at least of Elkind's (1961, 1962), and possibly Towler & Wheatley's (1971), judgements concerning the use of density arguments in distinguishing conservation from non-conservation answers.

With one possible exception, already referred to, all non-conservers of volume were conservers of weight and either used the customary arguments of identity, compensation or both to explain their assertions or commented on the independence of weight and shape.

Non-conservation of volume answers, for the moment ignoring the water displacement question which is discussed later, could be classified into three well-defined groups:

(1) A group of students who believed volume was dependent on shape. These made statements typified by a male arts graduate who in answer to both prediction and judgement questions said, 'No, the shapes have changed and although weight is the same volume will be different.'

(2) A group of students who believed volume was dependent on surface area. A typical answer to prediction and judgement questions for this group was, 'No, larger volume because of increase in surface area.'

(3) A small group of students who attempted to rely on formulae. One, a male science graduate, wrote in answer to the questions on prediction and judgement, 'No! The sausage-shaped object will have a larger volume, because of $\pi r^2 h$ (h depending on the length).' Another answer, provided by a female arts graduate, showed considerable confusion. She wrote, 'No (not necessarily). They only have the same volume if the volume of the cylinder $\pi r^2 l$ = volume of sphere. That is $r^2 l = \pi r^3$ (don't know the rest of the formula). That is if $r = l$ this is a unique case.'

As commented by Szeminska (1965), 'poorly assimilated information can become a mnemonic burden...'. (See also Rowell & Hoffmann, 1975.)

Of the 28 students (see Table 1) who gave non-conservation answers to questions concerning interior volume (see Piaget *et al.*, 1960) nine gave incorrect, but consistent, answers to the water displacement question (occupied volume), five failed either to give any distinctive answer – saying only that the water levels rose in each beaker but not commenting on the relative rise in level – or gave a correct answer but did not, or could not, explain it. Fourteen gave a correct answer together with an explanation but, with two exceptions, the explanation was incorrect. Twelve of the 14 believed the volume of water displaced was dependent on the weight of the object immersed, e.g. this answer, given by a male science graduate, sums up the position adopted by the group: 'The water level will be the same in each case because volume of water displaced depends upon the weight, and I have mentioned in question 1 that they weigh the same.' (One wonders to what extent a non-understanding of Archimedes' principle, half-remembered from high-school days, has contributed to this belief.) The remaining two explanations used correct wording but in at least one case it seems doubtful that the student fully understood the implications of her statements. In answer to the displacement question this student, a female arts graduate, wrote, 'If the water completely covers both pieces of plasticine the water level should rise an equal distance, providing that both pieces are of the same weight and density as was originally assumed.' Verbally a completely acceptable explanation of the situation. However, her answers to the prediction and judgement questions concerning what would happen to the volume of the plasticine when rolled into a sausage revealed her confusion over the interrelation of volume, weight and density. She wrote, 'If they are of equal density and weight but different size they should not have the same volume.'

DISCUSSION

The results of the present study provide some support for the generality of the findings of Elkind (1962) and Towler & Wheatley (1971), although the actual percentage of non-conservers reported here is substantially lower presumably, in part at least, because of the non-acceptance of density arguments as possibly indicative of conservation in the earlier studies, and the lack of suitable questions within the methodology employed which would allow any decision to be made on the issue. Nevertheless, a substantial proportion of people, even among the intellectual elite, either have no conception of what is meant by the term volume or, at best, have a very limited, and limiting, conception of it. In our sample of postgraduate Diploma in Education students 18·1 per cent were in this category with, as mentioned earlier, a significantly smaller proportion of females than males demonstrating an abstract conception of volume, irrespective of whether they were graduates in arts or the sciences.

Four problems are raised by these results and each is discussed briefly in the order shown. They are:

Question *a*. Why do people who demonstrate in some situations – by use, not reflexion – a knowledge of the logical processes of identity, reversibility and compensation, fail to apply them to the concept of volume?

Question *b*. Why is there a sex difference in this application?

Question *c*. Why is there no arts/science difference?

Question *d*. What are the educational implications?

Question a. Failure to conserve volume

The first problem is an extreme manifestation of the more general notion of *horizontal décalage* – a name given by Piaget to the observation that intra-individual differences exist in the ages at which problems are solved which obey identical structural laws but involve different content. For example, using a standard Piagetian approach, individuals invariably conserve amount before volume, and, as shown in the experiment described earlier, many never conserve volume.

What is peculiar to the volume problem which makes it difficult, particularly so for some people? There appears to be more than one factor contributing to an explanation. First, there is the fact that experiences relevant to the concept are complicated by the differential compressibility of substances. The fact that quantity remains constant in a given transformation need not mean that volume does. It is possible with all substances to change shape but not volume, it is possible with compressible substances also to change volume but not shape or alter both together. Possibly, at least partly for this sort of reason, volume and shape never become clearly differentiated in the minds of non-conservers, and such difficulties can only be enhanced by the fact that both are visually perceived three-dimensional attributes of a body. In this situation the mental operations of identity and reversibility provide no help in making a distinction. Even if Piaget is correct and 'identity' is the result of a combination of direct and inverse operations (a contention supported by the investigations of Smedslund, 1964) no separation of volume from shape (or surface area), with which it is confused, is *necessitated* solely by the operation of identity or the form of reversibility which describes negation. When a sausage of plasticine is rolled back into a ball it 'regains' its initial volume and regains its initial shape and surface area, and both of the latter *do* change when the plasticine is transformed from a ball into a sausage. All non-conservers of volume in the study reported here declared the volumes of the two pieces of plasticine would again be identical if the sausage were rolled back into a ball. It seems plausible to suggest that it is not until the internalized system of references required for 'reading' both volume and shape incorporates accurate estimation, or measurement, or the generalized knowledge resulting from these actions that the concepts are likely to become disentangled and emerge as independent abstractions.

Comments such as these provide a tentative explanation of the relative difficulty of extension of the conservation laws to volume and hence must contribute to any explanation of horizontal *décalages*. However, the personal nature of problem difficulty should not be overlooked. Undoubtedly individual differences exist in 'objectively describable' exposure to situations concerning volume in some way, but it seems unlikely that they can be held totally responsible for the high percentage of people, especially among the intellectual elite, science and arts graduates alike, who fail to achieve an abstract conception of the term. Differences between the individuals who give meaning to the experiences as *they* experience them, rather than just frequency and nature of the situations as they might be described by an observer seem of at least equal importance. And, if this is accepted, such differences also may contribute to the observed intra-individual horizontal *décalages*, and, indeed, to the establishment of conservation in any situation.

A piece of research particularly pertinent to the matter under discussion here is the study of individual differences in the apparent salience of various aspects of visually presented patterns reported by Taylor (1960), and discussed also by MacFarlane Smith (1964). Taylor gave a battery of tests, including the NIIP test of memory for designs and the Goodenough Draw-a-Man test, to 200 14-year-old children in an attempt to examine the nature of spatial ability and its possible relationship to attainment in geography. The illuminating feature of Taylor's work is that he found the same children achieved very different results on the tests depending on whether he marked for correction of proportion, defined as 'perception of the parts of the design in relation to the whole, i.e. a due relation in size between the parts and the whole had to be perceived' or correctness of detail, defined as 'the number of items or parts which were perceived as making up the composite whole' (Taylor, 1960). In view of what has been said earlier in this discussion, an individual who attends in a quantitative-analytic manner to the separable parts of a situation in relation to the whole seems more likely to be able to construct the conservation laws and progressively free them from their specific content than one who does not differentiate in this manner.

Although speculative, the differences in ways of experiencing a spatial stimulus noted by Taylor (1960) may find broader expression and usefulness in the notion of cognitive style. The possible importance of this variable has been emphasized by Pascual-Leone (1970) in his attempt to construct a performance model of Piagetian stage developmental theory. And, by implication at least, considerable stress seems to have been laid on an analytic-reflective style of encoding and processing information by several other authors discussing mechanisms leading to the construction and generalization of the conservation laws (see for example, Halford, 1970; Furby, 1972; Klahr & Wallace, 1973).

Question b. Sex differences

The second problem for discussion is the significant sex difference in performance on the volume conservation problem in spite of the availability of the operational structures underlying conservation in the competence repertoires of both groups as shown by their performance on the weight problem. Elkind (1962) attempted to explain the difference on the basis of a social hypothesis. He proposed that 'in adolescence social roles take over the direction of conceptualizing activity and that certain roles (academic, professional, scientific) were more conducive than others to attaining the volume conception'. The continued sex difference at first-year college level was fitted into the same explanatory system by maintaining that 'Many young women enter colleges or universities not because of genuine academic, professional or scientific roles but because of the social opportunities (the evidence for this is that a much higher percentage of women than men change schools and drop out of school before graduation)'. Social roles may provide a partial explanation of the difference in the study reported here too – the differential expectations for males and females which are held by society may be reflected in differential demands and opportunities provided for the two sexes to gain the experiences necessary for volume conservation; although some doubt is cast on this interpretation by the argument that if it is true it should be equally true for the conservation of other qualities of objects too, and there is no evidence from this study of a sex difference in conservation of weight. As all students in the sample were *post-graduate* teachers in training the non-adoption of academic, professional or scientific roles by the females seems less likely to be an important factor here than in the American studies.

On another level, a possible explanation of the difference is provided by the discussion of the first problem. If individuals differ in their cognitive style and this is associated with the production and generalization of the conservation laws, testable deductions are that more males than females are analytic, and secondly, in general it is the analysers who more readily achieve success in these problems. The first contention is strongly supported by the work of Taylor (1960), who found the performance gap between boys and girls on the NIIP spatial test, as marked according to the original scheme, widened markedly when scored for correctness of proportion. When marked for detail, girls scored rather higher than boys. In relation to his studies Witkin *et al.* (1962) comments, 'Women as a group tend toward a global approach in their perceptual and intellectual functioning, men towards an analytic approach'. The contention is supported also by the reported sex differences in ability to conserve (see, for example, Goldschmid, 1967) and, somewhat indirectly, by the numerous studies of sex differences, many of them summarized by Maccoby (1967), which indicate a general superiority of males in areas related to analytic ability, for example, mathematical reasoning, spatial ability and breaking set, and of females in areas related to global or synthetic functioning, for example, perceptual speed. The second contention is supported by the work of Pascual-Leone (1970) who, in his performance model of Piagetian theory, postulates cognitive style as a variable affecting the functional size of an age-dependent central computing space (M) and reports data indicating that 'field-dependent subjects frequently function with an M space inferior to their structural capacity'.

Question c. Graduates in arts and graduates in science

The third question concerns the lack of any significant difference between graduates in the arts and the sciences in their understanding of volume. *A priori* it might be expected that science- rather than arts-oriented individuals would have a better grasp of the meaning of a basic scientific term such as volume. The fact that this appears not to be the case, for this relatively small sample anyway, indicates that a scientific specialization has added nothing extra to common experience in this matter – no doubt such 'elementary knowledge' is expected of students by those teaching university courses in science. Could differences in cognitive style again aid in the provision of a parsimonious answer?

Question d. Some possible educational implications

In conclusion, a few comments may be made concerning the educational implications, both of the empirical study reported here and of the theoretical framework which has been used to provide a tentative explanation of it.

It has been argued by Kagan and his co-workers (1964) that cognitive style may be differentially effective in different academic disciplines, and by Satterly & Brimer (1971) that this line of reasoning may be extended to different functions within supposedly homogeneous content. What tentatively is suggested here is that the development of generalized structures of knowing may be in some way aided or hindered by the way in which an individual attempts to give meaning to a situation. Further research is needed to determine the viability of these remarks. If they are found useful a number of further matters arise: there is the question of establishing the peculiar strengths and weaknesses of the different styles, the extent to which they are alterable, and the nature and amount of teacher-directedness which may be necessary to aid individuals to achieve competence in different situations.

Secondly, on a more specific level, as pointed out by Lang (1972) in his analysis of the difficulty of a number of physical concepts, 'Designers of some [high school] curricula assume conservation of volume, e.g. Nuffield Physics (O level) and several other projects suggest using a monolayer of oil on water at the age of 11 to 12 years to provide an order of magnitude determination of the volume and hence the height of a molecule.' Judging from the results reported here, and those of Elkind (1961, 1962) and Towler & Wheatley (1971), this is a very rash assumption and teaching based on it is unlikely to result in any worthwhile learning for those children whose field of application of the relevant operative structures does not include volume. Before graduating to experiments assuming volume conservation children must first gain sufficient guided experiences involving the concept and emphasizing the operation of the structures 'necessitating' conservation – assuming these to be available in their competence repertoires. Presumably such experiences should include volume as space which can be occupied and space which is occupied, measurement (including 'filling in'), shape transformation (initially with non-compressible substances) followed by further measurement (and 'filling in') with particular emphasis on compensatory changes in the three dimensions – this being preceded by similar experimentation with area.

The third, and last point, is a note of concern. The results of tests reported here were gathered from a group of graduate teachers in training almost all of whom have now entered the teaching profession. How many non-conservers are teaching children about volume, and what are they teaching?

References

Elkind, D. (1961). Quantity conceptions in junior and senior high school students. *Child Dev.* **32**, 551–560.

Elkind, D. (1962). Quantity conceptions in college students. *J. soc. Psychol.* **57**, 459–465.

Furby, L. (1972). Cumulative learning and cognitive development: elaboration and implications of a pre-theoretical model. *Hum. Dev.* **15**, 265–286.

Goldschmid, M. L. (1967). Different types of conservation and non-conservation and their relation to age, sex, IQ, MA and vocabulary. *Child Dev.* **38**, 1229–1246.

Halford, G. S. (1970). A theory of the acquisition of conservation. *Psychol. Rev.* **77**, 302–316.

Kagan, J., Rosman, L., Day, D., Albert, J. & Phillips, W. (1964). Information processing in the child: significance of analytic and reflective attitudes. *Psychol. Monogr.* **78**, 1.

Klahr, D. & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cogn. Psychol.* **4**, 301–327.

Lang, W. A. F. (1972). How difficult are some concepts of physics. *Aust. Sci. Teach. J.* **18**, 59–63.

MacFarlane Smith, I. (1964). *Spatial Ability: its Educational and Social Significance.* London: University of London Press.

McKinnon, J. W. & Renner, J. W. (1971). Are colleges concerned with intellectual development? *Aust. J. Phys.* **39**, 1047–1052.

Maccoby, E. E. (ed.) (1967). *The Development of Sex Differences.* London: Tavistock Publications.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta psychol.* **32**, 301–345.

Piaget, J., Inhelder, B. & Szeminska, A. (1960). *The Child's Conception of Geometry.* London: Routledge & Kegan Paul.

Rowell, J. A. & Hoffmann, P. J. (1975). Group tests for distinguishing formal from concrete thinkers. *J. Res. Sci. Teach.* **12**, 157–164.

Satterly, D. J. & Brimer, M. A. (1971). Cognitive styles and school learning. *Br. J. educ. Psychol.* **41**, 294–303.

Smedslund, J. (1964). Concrete reasoning: a study of intellectual development. *Child Dev. Monogr.* **29**, 2.

Szeminska, A. (1965). The evolution of thought: some applications of research findings to educational practice. *Monogr. Soc. Res. Child Dev.* **30**, 47–57.

Taylor, C. (1960). A study of the nature of spatial ability and its relationship to attainment in geography. *Br. J. educ. Psychol.* **30**, 266–270.

Towler, J. O. & Wheatley, G. (1971). Conservation concepts in college students: a replication and critique. *J. gen. Psychol.* **118**, 265–270.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R. & Karp, S. A. (1962). *Psychological Differentiation: Studies of Development.* New York: Wiley.

(*Manuscript received* 4 *July* 1974; *revised manuscript received* 5 *December* 1974)

Br. J. Psychol. (1976), 67, 1, *pp.* 11–15
Printed in Great Britain

THE GOLDEN SECTION HYPOTHESIS

BY J. BENJAFIELD AND J. ADAMS-WEBBER
Department of Psychology, Brock University, St Catharines, Ontario, Canada

The golden section is a proportion the aesthetic properties of which have been extolled since antiquity. The data from five experiments in which subjects made dichotomous judgements of acquaintances on bipolar dimensions (e.g. *pleasant–unpleasant*) were reported. These data indicated that the mean proportion of positive adjectives used in making interpersonal judgements is an excellent approximation of the golden section. An explanation of this finding was offered in terms of Berlyne's 'strikingness hypothesis'. It was suggested that Boucher & Osgood's 'Pollyanna hypothesis' should be extended to include the possibility that, by tending to organize his judgements in the golden section ratio, the person is able to pay special attention to negative events.

The golden section is that point on a line which divides it into two segments such that the smaller is to the larger as the larger is to the whole line. As shown in Fig. 1, this means that the major segment (AC) must be approximately 0·62 of the line (AB). The golden section has had a ubiquitous influence on Western thought. The Pythagoreans took a mystical attitude toward it; Kepler called it the 'divine proportion'; and artists and architects have often incorporated it in their work (Huntley, 1970). From Fechner (1876) to Berlyne (1971), psychologists have done experimental work on its aesthetic properties. These experiments have tended to show that subjects find the golden section proportion more pleasing than any other.

The golden section may profitably be considered in conjunction with the Pythagorean notion that human experience can best be described in terms of pairs of opposites, such as *good v. evil, male v. female* and *straight v. curved* (Wheelwright, 1966). Plato, in the *Phaedo*, describes the Pythagorean belief that the body is maintained by a 'tensive relation' that exists between the pairs of opposites, and that the 'soul is a harmony' which is the result of the proper balance between opposites. The idea that successful living consists in finding the right proportion, or 'mix', of opposing tendencies is rooted in Pythagorean doctrine. As Russell (1945) points out, Pythagoras has influenced Western thought to a degree unmatched by anyone else. Thus it is of great interest that Osgood & Richards (1973) find ideas highly similar to the Pythagorean ones to be expressed in the ancient Chinese conceptions of Yang and Yin (Wilhelm, 1967).

Osgood & Richards' discussion of Yang and Yin leads them to the following conclusions: (1) cognition is organized along bipolar dimensions; (2) one of the poles of each dimension is psychologically positive (Yang) and the other negative (Yin); and (3) dimensions are related in parallel, positives with positives and negatives with negatives. Osgood & Richards point out, as does Deese (1973), that the positive member of a pair of opposites is usually the one which: (*a*) came into the language at an earlier date; (*b*) occurs more frequently in the language; and (*c*) is the first to be used correctly by children. The positive category (Yang) can be thought of as the original unity which becomes differentiated into a positive and a negative part.

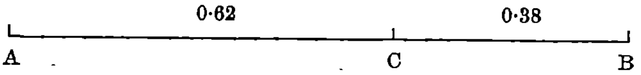


Fig. 1. The golden section.

A question which, in various forms, has attracted great interest throughout history is the following: What proportion should Yang and Yin bear to one another so as to yield the greatest harmony? We, of course, cannot answer such a transcendental question. However, a closely related empirical hypothesis suggests itself when this question is viewed from the perspective of the golden section. The hypothesis to be examined in this paper is that whenever subjects differentiate one thing into two, they tend to do so in a way that approximates the golden section. More specifically, we will present data indicating that when a group of subjects are given a task which requires them to divide a class of people into two classes on the basis of bipolar dimensions, the result is two classes which are in the golden section ratio. We will then show how a suggestion of Berlyne's (1971) concerning a possible reason for the aesthetic properties of the golden section can be generalized to account for this finding.

So formulated, the golden section hypothesis would predict that when subjects make dichotomous judgements about a series of acquaintances (figures) in terms of bipolar dimensions, they will tend to use positive adjectives 62 per cent of the time, and negative adjectives 38 per cent of the time. The data from five studies using Kelly's (1955, 1969) repertory grid task were used to evaluate this hypothesis.

In the first study, Kelly's (1955) original list of 19 role titles (e.g. Self, Mother, Father, Sister, etc.) were used to elicit the names of nineteen acquaintances from each subject individually. The name of each figure was recorded on a separate card. All 19 cards were shuffled thoroughly and handed to the subject, who was asked to sort them into two piles according to whether they were, for example, *strong* or *weak*. All figures designated as *strong* (and other positive terms) were scored '1' and those designated *weak* (negative) were scored '0' so that the results of subjects sorting could be recorded as a single 19-digit profile. (This score is similar to Bannister & Mair's (1968) 'maldistribution score'.) Then the subject was presented with the next dimension (e.g. *excitable–calm*), the cards were reshuffled, and he was asked to sort the figure cards again into two piles on the basis of that dimension. This procedure was repeated until the subject had sorted the same figures successively on 12 dimensions. Each dimension was written on a separate card. All 12 cards were shuffled thoroughly before they were presented to the subject The dimensions used were the same as those used by Warr & Coffman (1970): (1) *generous–stingy*;* (2) *pleasant–unpleasant*; (3) *true–false*; (4) *fair–unfair*; (5) *active–passive*; (6) *energetic–lethargic*; (7) *sharp–dull*; (8) *excitable–calm*; (9) *strong–weak*; (10) *bold–timid*; (11) *hard–soft*; and (12) *rugged–delicate*. Each of the three semantic differential components of connotative meaning (Osgood *et al.*, 1967) are represented by four dimensions: Evaluation (1–4); Activity (5–8); and Potency (9–12). For every dimension, the first adjective is positive, the second negative. Subjects were 30 Brock University undergraduate volunteers (15 men, 15 women).

---

* Warr & Coffman originally used *mean*. We substituted *stingy* for it because we found that Canadians had difficulty understanding *mean* as the opposite of *generous*. *Stingy* was used in all experiments reported in this paper except for the third (Adams-Webber & Benjafield, 1973), in which *mean* was used.

The total number of positive adjectives used in the grid task was calculated for each subject. The grand mean of these scores, expressed as the proportion of positive adjectives used to the total number of adjectives used, was 0·62 (S.D. = 0·07). This result means that subjects tend to allot figures to the positive poles of dimensions 62 per cent of the time, and to the negative poles 38 per cent of the time.

The procedure described above was repeated using 12 figures selected from Kelly's original list, plus six others. These last six consisted of three evaluatively positive figures (*warmest person; most active person; strongest person*) and three evaluatively negative figures (*coldest person; most passive person; weakest person*). Subjects were 30 Brock University undergraduate volunteers (15 men, 15 women). The grand mean number of positive adjectives used, expressed as the proportion of positive adjectives used, was 0·62 (S.D. = 0·07) for the first 12 figures, and 0·62 (S.D. = 0·07) for the last six.

The data from two previously published studies (Adams-Webber & Benjafield, 1973; Benjafield & Adams-Webber, 1975) were reanalysed. These studies followed an identical procedure to that described above, except that only the 12 figures selected from Kelly's original list were used. The grand means of the scores derived from these data, again expressed as the proportion of positive adjectives used, were 0·63 (S.D. = 0·08) and 0·61 (S.D. = 0·08), respectively.

Finally, repertory grids containing the same 12 figures and 12 dimensions were administered during class time to a group of 62 students in an introductory psychology course at Brock University. Each student received a grid with the 12 role-titles across the top and the 12 dimensions listed down the right-hand side. The order of the dimensions was arbitrary and the same for each subject. The left-hand member of each pair was positive half the time, and negative half the time. The subject filled in the names of figures which correspond to the 12 role-titles. He then put a check mark for each figure that was best described by the left-hand member of each adjective pair, and left a blank if the figure was best described by the right-hand member. The grand mean number of positive adjectives used, again expressed as a proportion, was 0·63 (S.D. = 0·06). Thus all six results are within 0·01 of the golden section. Not only were there no significant sex differences in any of these experiments, but also, the means for men and women were almost identical in each case.

DISCUSSION

Berlyne (1971, p. 232) suggests that the popularity of the golden section may be due to the fact that 'it allows the minor element to occupy a proportion of the whole that makes it maximally striking'. This idea derives from the work of Frank (1959, 1964), who showed that when subjects are asked to arrange colours so as to make one colour stand out against the others, that colour is used between 0·37 and 0·38 of the time. In other words, in order to make a figure maximally striking, the proportion of figure to background is very close to the golden section. As Berlyne & Frank have pointed out, there are good information-theoretic reasons why this should be so. Average information is given by summing the quantities $-p(x) \log p(x)$, where $p(x)$ is the probability of occurrence of each category (Garner, 1962, pp. 20–4). The quantity $-p(x) \log p(x)$ is a maximum when $p(x)$ is equal to approximately 0·37. That is, the contribution a particular category of events makes to average information is a maximum when the relative frequency of that category is equal to approximately 0·37.

The implication of our repertory grid findings would then be that subjects tend to organize their interpersonal judgements along bipolar dimensions so as to make negative events maximally striking. The person tends to construe others in such a way that negative events, taken as a whole, 'stand out' as 'figure' against a positive 'background'. Negative adjectives label 'marked', or deviant, events (Deese, 1973). Deviant events are, almost by definition, precisely those which are problematic, or troublesome, for the person. They 'bear watching'. Thus an information-processing strategy which, at an abstract level, made negative events maximally striking, would make considerable adaptive sense. By arranging his judgements in the golden section ratio, the person is able to pay special attention to negative events considered as a whole.

The golden section hypothesis goes considerably beyond Boucher & Osgood's (1969) Pollyanna hypothesis. The latter holds that 'there is a universal human tendency to communicate about the positive aspects of life' (Osgood & Richards, 1973, p. 410). The Pollyanna hypothesis is correct as far as it goes, but it is also partially misleading because it implies, as Deese (1973) points out, that 'We are all natural optimists'. The golden section hypothesis suggests that, while we construe most events positively, we attempt to create a harmony between positive and negative events such that the latter make a maximal contribution to the whole.*

It is tempting to go further. Perhaps the golden section can be seen as representing the kind of 'union of opposites' which Jung (1968) has described as optimal for the person. Perhaps there is an explanation for the aesthetic quality of the golden section in our results. Since our interpersonal judgements tend to be organized in this ratio, might not objects please us to the extent that they can support a 'projection' of our own cognitive structure? But such speculation is beyond the scope of this paper.

This research was supported by Canada Council Grant S74-0686. Discussions with Linda Bebout, Harry Hunt, David Jordan and Sid Segalowitz were very helpful. Paula Harder and Antoon Leenaars helped gather some of the data.

---

* Data supporting the golden section hypothesis may also be found in the Osgood & Richards experiment itself. They gave subjects 200 sentences of the form X is (adjective) —— (adjective), such as X is *dangerous* —— *empty*. The subject's task was to decide whether *and* or *but* should be used to fill the blank. From the viewpoint of the golden section hypothesis, the subject is being asked to divide the sentences into two classes: those which take *and*, and those which take *but*. Since *and* is positive and *but* negative (Osgood & Richards, 1973, p. 397), subjects should use *and* 62 per cent of the time, and *but* 38 per cent of the time. Osgood and Richards (1973, pp. 401, 409) report that *and* is used, on the average, 62·5 per cent of the time and *but* 37·5 per cent of the time – impressively close to the golden section ratio.

References

Adams-Webber, J. & Benjafield, J. (1973). The relation between lexical marking and rating extremity in interpersonal judgement. *Can. J. behav. Sci.* **5**, 234–241.

Bannister, D. & Mair, J. (1968). *The Evaluation of Personal Constructs.* London: Academic Press.

Benjafield, J. & Adams-Webber, J. (1975). Assimilative projection and construct balance in the repertory grid. *Br. J. Psychol.* **66**, 169–173.

Berlyne, D. E. (1971). *Aesthetics and Psychobiology.* New York: Appleton-Century-Crofts.

Boucher, J. & Osgood, C. E. (1968). The Pollyanna hypothesis. *J. verb. Learn. verb. Behav.* **8**, 1–8.

Deese, J. (1973). Cognitive structure and affect in language. In P. Pliner, L. Krames & T. Alloway (eds.), *Communication and Affect.* London: Academic Press.

Fechner, G. T. (1876). *Vorschule der Ästhetik.* Leipzig: Breitkopf & Härtel.

Frank, H. (1959). *Grundlagenprobleme der Informationsästhetik und erste Anwendung auf die mime pure.* Quickborn: Schnelle.

Frank, H. (1964). *Kybernetische Analysen subjektiver Sachverhalte.* Quickborn: Schnelle.

Garner, W. R. (1962). *Uncertainty and Structure as Psychological Concepts.* New York: Wiley.

Huntley, E. H. (1970). *The Divine Proportion.* New York: Dover.

Jung, C. G. (1968). *Psychology and Alchemy.* Princeton, N.J.: Princeton University Press.

Kelly, G. A. (1955). *The Psychology of Personal Constructs.* New York: Norton.

Kelly, G. A. (1969). A mathematical approach to psychology. In B. A. Maher (ed.), *Clinical Psychology and Personality.* New York: Wiley.

Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. (1967). *The Measurement of Meaning.* Urbana, Ill.: University of Illinois Press.

Osgood, C. E. & Richards, M. M. (1973). From Yang and Yin to *and* or *but*. *Language* **49**, 380–412.

Russell, B. (1945). *A History of Western Philosophy.* New York: Simon & Schuster.

Warr, P. B. & Coffman, T. L. (1970). Personality, involvement and extremity of judgement. *Br. J. soc. clin. Psychol.* **9**, 108–121.

Wheelwright, P. (1966). *The Presocratics.* New York: Odyssey Press.

Wilhelm, P. (1967). *The I Ching or Book of Changes.* Princeton, N.J.: Princeton University Press.

(*Manuscript received* 14 *January* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 17–22
Printed in Great Britain

INSPECTION TIME AND MEASURED INTELLIGENCE

By T. NETTELBECK and M. LALLY
Department of Psychology, University of Adelaide, Adelaide, South Australia

Ten subjects with WAIS Full Scale IQ scores ranging from 119 to 47 were required to discriminate between two lines of markedly different length, exposed in random order for ten different durations, by pressing one of two keys. Estimates of inspection time λ, the rate at which sensory input is accumulated and passed to subsequent decision processes, were calculated directly from the psychometric functions obtained. λ was found to correlate negatively with subtests contributing to Performance IQ. There was a substantial degree of consistency between estimates of λ on two different occasions and training did not appear to affect the index, although differences between the two estimates increased as the value of the first estimate increased. Mean overall response latency did not correlate significantly with intelligence and measures obtained from retarded subjects were very like those reported in similar experimental situations utilizing normal subjects. This unexpected finding is interpreted as suggesting that, in this situation, retarded subjects have been prepared to respond on the basis of less evidence than is usually required by intellectually normal subjects.

Vickers *et al.* (1972) have outlined a number of measures of perceptual performance developed recently within a common theoretical framework. They propose that the general features of simple discriminative judgements may be accounted for by a model, according to which it is assumed that, before making a decision, the subject takes a series of observations of sensory input at a steady rate, accumulating the evidence obtained until the increment favouring one alternative reaches a predetermined criterion. The subject's behaviour is therefore seen as conforming to an optional stopping process.

A number of optional stopping models have been proposed, minor differences between the detailed properties of the various models stemming mainly from what constitutes the criterion for decision. However, all assume that the subject's criterion reflects the level of caution with which a response is made. This type of model therefore provides an account of the speed–accuracy trade off which has been found to occur in a wide variety of perceptual tasks. A preferential bias towards a particular response is accounted for by the assumption that the subject may adopt a lower criterion for that response, and therefore be prepared to make it on the basis of less evidence.

All theoretical models of this kind predict that the level of noise in the perceptual system will limit the subject's sensitivity, and that the pattern of correct responses, in a two-choice task in which discriminability is varied, will approximate the form of a cumulative normal ogive. This psychometric function will become steeper as noise is reduced and sensitivity increased, or as the level of caution is increased.

Vickers *et al.* have suggested that, if obtained under specified conditions, the psychometric function can be used to obtain a measure of the rate at which the subject samples the sensory input. They term this index 'inspection time' λ, proposing that an estimate of λ is provided by the stimulus duration at which performance in an extremely easy discrimination task is virtually free from error. Vickers *et al.* have outlined a psychophysical procedure which constitutes an incomplete method of constant stimulus differences, in which the proportions of the subject's correct responses at various durations of stimulus exposure are used to derive the upper half of the psychometric function. From this function λ may be calculated directly, using the 97·5 per cent level of confidence as a criterion for virtually error-free performance.

Since λ is obtained from the probability of a correct response at various stimulus exposure durations, and under conditions which prevent the accumulation of data, it is free to vary independently of response latency. Results from two experiments involving university undergraduates suggest that λ is independent of response latency, and that 100 msec. is a good estimate of the average time needed to make one inspection of the sensory input (Vickers *et al.*, 1972).

In the investigation described here, the relationship between individual differences of intelligence, as measured by the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1958), and inspection time λ has been examined. This follows a proposal from Savage (1970) that differences in intellectual ability may be attributable in part to the speed with which 'some kind of search or scanning mechanisms' operate (p. 37). O'Connor & Hermelin (1965) have made a similar suggestion, arguing that the restricted input of sensory data may be as significant a factor in retardation as short-term memory deficit.

METHOD

Apparatus

Stimuli drawn on cards were presented tachistoscopically. A small luminous spot, placed in the area where the stimulus difference would appear, provided an attentional cue. Stimulus exposure activated an electronic timer. A response was made by pressing one of two morse keys, the latency being registered with 1 msec. accuracy.

Stimuli

Stimuli were similar to those described by Vickers *et al.* (1972), consisting of two vertical lines 24 and 34 mm long, and 10 mm apart, with their upper ends terminated by a third horizontal line. The display was viewed from a distance of 38 cm, so that the stimulus difference of 10 mm corresponded to a visual angle of 1·6°. In the absence of any evidence concerning the relationship between intellectual ability and the level of noise in the perceptual system, this figure was chosen, being twice that used by Vickers *et al.*

A trial began with the exposure of the stimulus display. A similar figure then overlaid the display with lines 44 mm long, making the stimuli indiscriminable and acting as a backward mask which remained exposed for 5 sec. Lines in the backward mask were made thicker than those in the stimulus figure, in order to offset small inconsistencies when centring the stimulus cards in the tachistoscope. The intertrial interval was approximately 4 sec., a sufficiently long period to allow the subject to recover from any inhibitory effects resulting from the previous stimulus and response (Baumeister & Kellas, 1968).

Stimuli were exposed at 10 different durations, with an equal number at each duration randomly presented throughout each successive block of 20 trials. Within each block, the occurrence of the shorter line on the left or right was equiprobable and the sequence random. Exposure durations of 10, 30, 50, 60, 70, 80, 90, 100, 115 and 130 msec. were used for nine subjects; with one subject durations were 60, 70, 80, 90, 100, 120, 130, 150, 200 and 300 msec. for reasons outlined below.

Subjects

Ten male subjects aged 16–22 years, and employed at a vocational rehabilitation centre, were selected on the basis of WAIS Full Scale IQ scores. Individual scores were 119, 108, 98, 85, 81, 80, 77, 76, 56 and 47. The three highest scorers had been admitted to the centre because of light injury or for assistance with behavioural problems. The remaining seven had been referred as being unlikely at the time to cope with open employment, and worked under sheltered conditions.

No subject had any physical disabilities likely to affect responding, and no drug treatment was involved. Monocular visual acuity was checked beforehand, using a different Snellen chart for each eye. All subjects volunteered to participate, and remained naive with respect to the experimental aims.

Design

Subjects attended individual sessions at the same time each day, completing five runs of 105 trials each. The first five trials of each run were for practice only and therefore excluded from analysis. As described in detail by Vickers *et al.* (1972), a measure of inspection time λ was calculated directly from the cumulative normal ogive with a mean of zero best describing the pattern of response probabilities obtained from each subject.

On the basis of measured intelligence, the 10 subjects were divided into two groups matched as evenly as possible. Five subjects were assigned to an experimental group in which they were given an additional two practice sessions of 100 trials each, during which the subject was told, following every trial, whether his response was correct or incorrect. The five control subjects did not attend these sessions and remained unaware that their experimental counterparts had been given special training.

In order to obtain a second measure of λ, all subjects completed a further six runs of 105 trials each. The initial run and the first five trials of each subsequent run were for practice only and were excluded from analysis.

Procedure

Before commencing the experiment, special care was taken to ensure that each subject understood the nature of the task and the use of the response keys. During initial training, long exposures ranging from 300 msec. up to 2 sec. were used. The importance of the accuracy of the response, rather than the speed, was emphasized during early training, and the subject was encouraged to respond both verbally and manually, to offset difficulty in establishing an association between the verbal instruction and the manipulation of the task (O'Connor, 1965). Subsequently the exposure duration was reduced gradually, the criterion accepted for learning being 12 successive correct responses, using an exposure duration of 130 msec. Nine subjects reached criterion within 40 min. of training. The tenth (Full Scale IQ = 47) was unable to learn the task within an hour, and it was only with considerable difficulty that this subject was taught to produce an error-free performance using an exposure duration of 300 msec. The range of 60–300 msec. was used for this person during subsequent experimental sessions.

The subject was directed to indicate the shorter line in the stimulus display by pressing the appropriate key as quickly as possible. It was emphasized that, once overwritten by the backward mask, the display would provide no further evidence.

The experimenter remained present throughout all sessions, encouraged the subject to pay attention before each trial, and continuously monitored data to detect any tendency for the subject to develop a preferred response (Gerjuoy & Winters, 1968). In fact, although a problem during initial training, no subject did this during experimental sessions.

RESULTS AND DISCUSSION

Inspection time λ

Examination of the data suggested a non-linear relationship between inspection time and Performance IQ. However, although the Spearman rank correlation coefficient has been adopted as more appropriate, using the Pearson *r* does not affect the outcome substantially.

Both the initial estimate of inspection time (λ_1) and the second estimate (λ_2)

Table 1. *Spearman rank correlations found between standard scores for subtests in the WAIS from which Performance IQ is derived, and two estimates of inspection time* (λ_1 *and* λ_2)

(The average correlation between the subtest concerned and each of the other subtests is also shown.)

	Digit symbol	Picture completion	Block design	Picture arrangement	Object assembly	Performance IQ
Inspection time: λ_1	−0·72*	−0·75*	−0·87**	−0·60	−0·86**	−0·92**
Inspection time: λ_2	−0·71*	−0·72*	−0·91**	−0·64*	−0·77**	−0·89**
Other Performance IQ subtests (average correlation)	+0·78**	+0·82**	+0·77**	+0·75*	+0·76**	—

* $P < 0\cdot05$; ** $P <$ 0 01 (two-tailed).

were highly correlated with Performance IQ as measured by the WAIS. It should be noted, however, that these correlations are likely to be overestimates of those in the general population, since the relatively small sample was selected to be representative of a wide range of measured intelligence. For λ_1, Spearman $\rho = -0\cdot92$ ($P < 0\cdot01$; two-tailed), while for λ_2, Spearman $\rho = -0\cdot89$ ($P < 0\cdot01$; two-tailed). These correlations and those with the separate subtests are shown in Table 1. It can be seen that the correlations between the estimates of inspection time and the various subtests were comparable in size with those indicating the internal consistency of the WAIS. The relationship with Verbal IQ was not significant, either for λ_1 (Spearman $\rho = -0\cdot32$; $P = 0\cdot37$; two-tailed) or for λ_2 (Spearman $\rho = -0\cdot41$; $P = 0\cdot26$; two-tailed).

Estimates of λ ranged from 98 to 554 msec. for the first measure, and from 100 to 804 msec. for the second. Five of the six measures obtained from the three subjects with IQ scores above 85 were within the range of 74–169 msec. found by Vickers *et al.* (1972), whereas 10 of the 14 estimates obtained from the retarded subjects fell beyond it.

Differences between λ_1 and λ_2, disregarding sign, increased as the value of the first estimate increased (Pearson $r = 0\cdot89$; d.f. $= 8$; $P = 0\cdot001$; two-tailed). However, despite this variability there was a substantial degree of consistency between the two sets of inspection times (Pearson $r = 0\cdot96$; d.f. $= 8$; $P < 0\cdot001$; two-tailed), with no more than chance variations between them ($F < 1\cdot0$). The effects of training were examined by analysis of covariance, in order to control for variability due to IQ. The main effects and the interaction between groups and the first and second determinations of λ were small and not significant ($F < 1\cdot0$ in each case), suggesting that training prior to obtaining the second estimate of λ did not influence the outcome, and that the measure of λ was little affected by practice.

Latency

Mean overall latency L was obtained during sessions concerned with both the first and second estimates of λ, by averaging the latencies for all responses made by a particular individual. Values ranged from 291 to 649 msec. for the first occasion and from 348 to 667 msec. for the second. As shown in Table 2, there was little evidence to suggest a relationship between either L_1 or L_2 and factors contributing to intelligence, correlations being small and not significant. This outcome was

Table 2. *Spearman rank correlations between two measures of mean overall latency (L_1 and L_2) and measured intelligence*

(Critical level ($P = 0{\cdot}05$; two-tailed); $\rho = 0{\cdot}65$.)

	Verbal IQ	Performance IQ	Full scale IQ
Mean overall latency: L_1	$-0{\cdot}07$	$-0{\cdot}25$	$-0{\cdot}15$
Mean overall latency: L_2	$+0{\cdot}33$	$+0{\cdot}16$	$+0{\cdot}18$

confirmed when the relationships between intelligence, and the average latencies at each of the 10 different exposure levels used to determine λ_1 and λ_2, were examined separately.

The absence of any substantial relationship between measured intelligence and response latency was unexpected, in view of the widely reported correlations between these two variables (Baumeister & Kellas, 1968). Furthermore, the values of mean overall latency L found in this experiment resembled closely those obtained by Nettelbeck (1973) with normal subjects, in two experiments carried out under conditions similar to those described here. Using 10 university undergraduates aged from 18 to 25 years, he found that measures of L ranged from 387 to 761 msec.

Reporting this work, Vickers *et al.* (1972) have pointed to the stability of L as a descriptor of individual performance in similar situations, and have proposed that the index provides some indication of the level of caution with which the subject responds. These authors have argued that increasing the number of inspections made of the stimulus display increases the time taken to make a response, but would normally enable the subject to accumulate additional evidence upon which to base a decision. This is obviously so when the display is present during the whole time required to arrive at a decision, but is also possible when it is not, in the sense that some data will enter the stores from random noise alone. Results from Vickers *et al.* suggest that the subject 'tops up' his stores to criterion by this means, and does not abandon an optional stopping strategy.

In this experiment, as in those reported by Vickers *et al.* (1972), there was substantial agreement between the measures of L obtained in the two sessions (Pearson $r = 0{\cdot}74$; d.f. $= 8$; $P = 0{\cdot}01$; two-tailed). At the same time, L was not significantly related to λ_1 (Pearson $r = 0{\cdot}25$; d.f. $= 8$; $P = 0{\cdot}50$; two-tailed) or to λ_2 (Pearson $r = -0{\cdot}42$; d.f. $= 8$; $P = 0{\cdot}23$; two-tailed).

In summary, then, mean overall latency L was not related to measured intelligence despite the substantially longer period of time required by the retarded for the accumulation of sensory data, as indicated by estimates of λ. Values of L appeared comparable with those obtained on a different occasion but under similar experimental conditions from intellectually normal subjects of the same age. Taken together, these findings suggest that the retarded subjects in this experiment were prepared to reach a decision on the basis of less evidence than was required by the intellectually normal subjects reported by Nettelbeck (1973). In other words, it appears that here intellectually retarded subjects may have adopted lower criteria for responding.

P2230

CONCLUSIONS

Although there are clear differences between the tasks presented in the various Performance subtests of the WAIS, these subtests have correlated consistently with inspection time. However, all subtests share in common a dependence upon the testee's ability to detect detailed relationships within the material presented, and to make appropriate selections from among competing stimuli, within a limited period of time. Results in this experiment suggest that performance on tests of this kind may be affected adversely by a slower rate of sampling sensory data.

Since the capacities measured by indices of the kind outlined here seem likely to be important in many forms of sensorimotor performance, it is felt that the procedures described might be developed to provide a useful adjunct to existing tests of measured intelligence. The suggestion that even mild degrees of retardation may be characterized by slower inspection time provides a possible explanation for the finding that retardates require longer periods of time in order to experience succession (Thor, 1970).

The measure λ is concerned directly with the assessment of fundamental discriminatory capacity, and any marked deviation from normal amongst the mentally retarded must have important implications for training methods and working conditions designed with a view to vocational habilitation. Whereas some training procedures might successfully change the level of caution with which a response is made, special planning would be necessary to accommodate basic limitations of information-processing capacity as measured by λ.

M. Lally received a studentship from APEX Foundation for Research into Mental Retardation Ltd. Thanks are due to the management and staff of Bedford Industries Vocational Rehabilitation Association Incorporated, South Australia, for their help and cooperation during the project, and to Professor A. T. Welford, Dr D. Vickers, Dr N. H. Kirby and Mr E. E. Rump for helpful comments on an earlier draft of this paper. A modified version of the paper was presented at a Conference in Experimental Psychology, Monash University, Victoria, 28–30 June 1974.

REFERENCES

BAUMEISTER, A. A. & KELLAS, G. (1968). Reaction time and mental retardation. In N. R. Ellis (ed.), *International Review of Research in Mental Retardation*, vol. 3. New York: Academic Press.

GERJUOY, I. R. & WINTERS, J. (1968). Development of lateral and choice sequence preferences. In N. R. Ellis (ed.), *International Review of Research in Mental Retardation*, vol. 3. New York: Academic Press.

NETTELBECK, T. (1973). An experimental investigation of some parameters affecting individual differences in perception. (Unpublished Ph.D. thesis, University of Adelaide.)

O'CONNOR, N. (1965). Learning and mental defect. In A. M. Clarke & A. D. B. Clarke (eds.), *Mental Deficiency: the Changing Outlook*, 2nd ed. London: Methuen.

O'CONNOR, N. & HERMELIN, B. (1965). Input restriction and immediate memory decay in normal and subnormal children. *Q. Jl exp. Psychol.* **17**, 323–328.

SAVAGE, R. D. (1970). Intellectual assessment. In P. Mittler (ed.), *The Psychological Assessment of Mental and Physical Handicaps*. London: Methuen.

THOR, D. H. (1970). Discrimination of succession in visual masking by retarded and normal children. *J. exp. Psychol.* **83**, 380–384.

VICKERS, D., NETTELBECK, T. & WILLSON, R. J. (1972). Perceptual indices of performance: the measurement of 'inspection time' and 'noise' in the visual system. *Perception* **1**, 263–295.

WECHSLER, D. (1958). *The Measurement and Appraisal of Adult Intelligence*, 4th ed. Baltimore: Williams & Wilkins.

(*Manuscript received* 4 *July* 1974; *revised manuscript received* 20 *February* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 23–29
Printed in Great Britain

IMPLICIT-SHAPE CONSTANCY AS A FACTOR IN PICTORIAL PERCEPTION

BY J. B. DERĘGOWSKI
King's College, University of Aberdeen

Schoolboys from the Ivory Coast and from Kenya were tested on a simple task intended to detect the effect of implicit-shape constancy. The presence of such an effect was confirmed. Comparison of two Ivorean subsamples differing in their responses to a request to build a model of a depicted structure showed that those subjects who built flat (2D) models were less influenced by the implicit-shape constancy than those subjects who built three-dimensional structures. This confirms a postulated link between the implicit constancy and perception of pictorial depth.

Difficulties in perception of pictorial depth have been demonstrated in several cultures, although by accident rather than by design most of the relevant work has been carried out in Africa. The problem has been discussed in four recent reviews: by Hudson (1967), who in his earlier papers (Hudson, 1960, 1962) brought the issue to the attention of psychologists; by Deręgowski (1971, 1973); and by Miller (1973). One can probably accept, in view of the evidence presented, that the difficulty exists. This brings forth a number of questions of interest to psychology at large. One of these concerns the mutual influence of various elements of a picture.

Deręgowski (1968) remarked that it might be that those of his subjects who responded to drawings as if they did not carry the implications of pictorial depth did not fail to see depth because they failed to recognize depth cues, but rather because, whilst looking at pictures, they restructured them by allocating lesser importance to depth cues than to other elements. Pictorial depth can be conveyed in a variety of ways. Blakemore (1973) lists seven devices by which painters create an illusion of depth on flat surfaces. There are, however, some configurations which do not appear to fall easily into categories listed by him. The Necker cube is a case in point. It does not involve linear perspective. In one of its perceived orientations it exhibits, admittedly, the effect which Blakemore terms 'position in the field', whereby an object (in this case the face of the cube), which is seen as being further away, is depicted higher than a 'nearer' object. In its 'reversed' orientation, however, the Necker cube is seen with its 'nearer' face higher up and hence contradicts all tenets of depiction, and yet, in spite of this, creates a remarkably strong illusion of depth. This 'Necker' depth effect might indeed have been responsible for Deręgowski's (1968) observation that subjects who failed to perceive pictorial depth in the meaningful pictures which Hudson's test constitutes did appear to perceive it when confronted with figures which can perhaps be described as vestigial cubes.

Gregory (1966, p. 173) offers a good illustration of a phenomenon which can perhaps be conveniently termed 'implicit-shape constancy'. It shows a boy in a street playing with a hoop. The hoop is shown as inclined and is therefore drawn as an ellipse. Yet the picture of the hoop is perceived as that of a usual circular hoop. The bulk of the evidence about the 'true' shape has to be derived from the ambient cues, not from the geometrical form drawn. The boy is likely under the circumstances depicted to hold a hoop. The 'true' shape of a depicted figure is thus implied by its pictorial surroundings. Implicit-shape constancy should also be observable when the implying elements are geometrical. Furthermore, in the case of such figures as cubes which have great power to evoke perception of depth, implicit-shape constancy should be observable even within those populations which exhibit difficulties in perception of pictorial depth.

The present paper is primarily concerned with investigation of the influence of this particularly strong depth cue upon perception of an implicit constancy of geometric patterns.

Method

Experimental plan

Two experiments were conducted: the first was intended to establish whether the 'implicit-shape constancy' was present in the population tested, the second to investigate the relationship between the implicit-shape constancy and perception of pictorial depth.

Experiment I

Hypothesis

That the subjects drawn from populations known to have difficulties with pictorial depth perception will show evidence of perception of implicit-shape constancy when confronted with strong implying cues.

Subjects

In order to increase the generality of the sample, the subjects were drawn from two populations: the Bukusu of Kenya and the Bualé of the Ivory Coast.

The Bukusu live in Western Kenya near Mount Elgon and belong to the Bantu group of tribes. The sample used was drawn from a *Harambee* secondary school in a small town. This school, in common with other *Harambee* schools, provides education to those pupils who have failed, generally owing to poor performance in examinations, to gain entry to Government secondary schools.

The Bualé live in Central Ivory Coast. The primary school from which the sample was drawn is in a large village in which a Swiss research centre has an established base. The Bualé belong to the West African group of tribes. Guerry (1970) provides an ethnographic description of the tribe.

The sample characteristics were as follows: *Bualé* ($n = 25$). Mean age 13·1 years, S.D. 1·2 years; mean years of schooling 6·8, S.D. 1·2. *Bukusu* ($n = 28$). Mean age 17·4 years, S.D. 2·1 years; mean years of schooling 8·2, S.D. 0·7.

The Bukusu sample was randomly subdivided into two equal groups, one of which participated in the main experiment and the other served as a control group.

Procedure

Essentially the same testing procedure was used with both samples, the only major difference being the language in which the subjects were instructed. In Kenya the instructions were given in English by the experimenter. In the Ivory Coast the instructions were given in French by a European co-worker whilst the experimenter merely recorded the results.

All testing was done on school premises either in empty classrooms or alfresco within the school grounds.

Subjects were tested individually. It was carefully explained to each subject that the experimenter was interested in learning how people see pictures. Each subject was told that he would be shown a drawing, that this drawing would then be taken away and he would have to choose from a set of drawings a drawing similar to that which he had just seen. The picture was exposed in the horizontal plane on the table immediately in front of the subject. After about 5 sec. the stimulus picture was removed and, after a lapse of about 3 sec., a card bearing five response choices was put in its place.

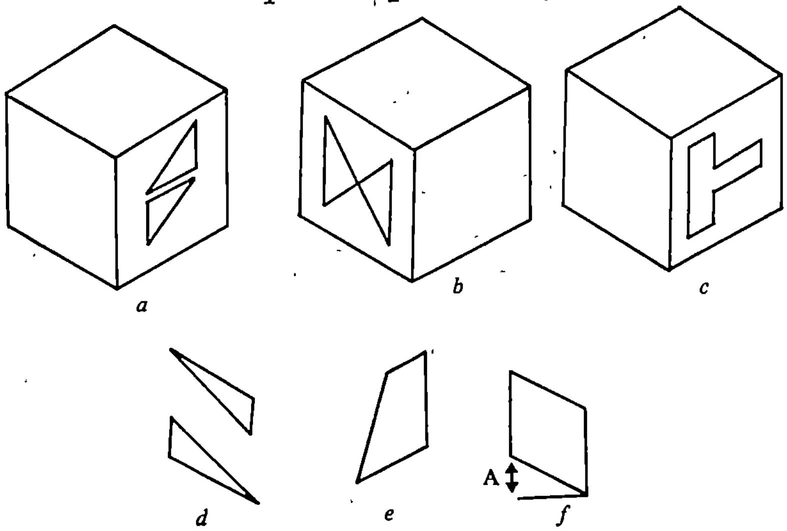


Fig. 1. Stimulus drawings. (*a–c*) Three of the six patterns used 'drawn on sides of skeleton cubes'. (*d–f*) The other three patterns. Each pattern was drawn both on its own and 'on a cube'. Two sets of six stimuli were thus obtained.

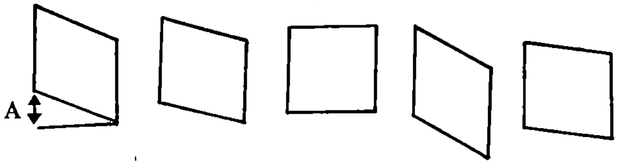


Fig. 2. Response drawings – main set. Six variants of one of the stimulus drawings differing in magnitude of angle A, and having all the lines which formed angle A = 0°, 7·5°, 15°, 22·5° and 30°.

The subject was told to choose a figure like the one he had just seen. The choice made was recorded, the card removed and the next stimulus figure presented. Each subject responded to six 'cube' and six 'plane' stimuli, the stimuli being presented in a random sequence.

Materials

Two sets of stimulus drawings, each consisting of six figures and two sets of response-choice figures, were used. Examples of the stimulus drawings are shown in Fig. 1. Each stimulus drawing included a simple geometric pattern consisting of either one or two simple figures, which the subjects had to recall. Three of these geometric patterns are shown in Figs. 1*d*, *e* and *f*. Using these patterns two types of stimuli were prepared. In one of these the geometric patterns were drawn on the face of an isometric skeleton cube (as shown in Figs 1*a*, *b* and *c*, wherein the remaining three geometric patterns are portrayed). In the second type the same patterns were drawn on plain pieces of card. Identical white cards (10 × 10 cm) were used for both 'cube' and plane stimuli.

Two response-choice sets were made up as follows. The experimental set consisted of six cards (10 × 28 cm) on each of which five drawings were drawn. One of these was identical with one of the geometrical patterns used and the other four were derivatives of this very pattern. They were derived by systematically varying the angle A (see Fig. 2) giving it one of the following values 0°, 7·5°, 15°, 22·5°, keeping the length of all the lines, which in the original pattern formed an angle A of either 30° or 90°, constant. The angle A was defined using skeleton cubes as the angle between a line in the plane of the paper perpendicular to the 'edge' of the depicted cube 'nearest' to the viewer, and the line forming a part of the depicted pattern and parallel to the 'non vertical edges' of the face of the cube on which the pattern was drawn.

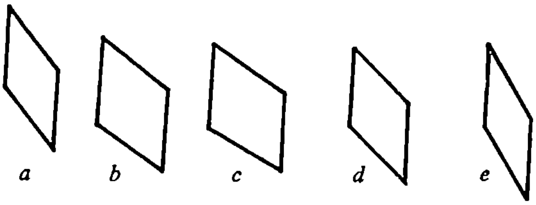


Fig. 3. Response drawing – control set. Six variants of one of the stimulus drawings. Angle A = 30°, 37·5°, 45°, 52·5° and 60°.

The figures on each response drawing were arranged in random order.

The control set was in all respects, but for the method of derivation of figures, identical with the experimental set. It differed in that the derivative drawings had angle A values of 37·5°, 45°, 52·5° and 60° (see Fig. 3).

Results

Each subject's total score was derived as follows. The responses he made to the 'plane' and 'cube' stimuli incorporating the same pattern were compared. When these responses did not differ, the same choice having been made under both circumstances the subject was awarded a score of '0'. When the stimuli chosen did differ the scores awarded were: '1' when the 'plane' stimulus led to a response showing greater effect of the implicit-shape constancy, and '-1' when the 'cube' stimulus did so. Since each subject responded to six stimuli his score could range between -6 and $+6$.

This method of scoring was adopted because the psychophysical values of depicted cubes do not appear to be easily definable (Ekman & Junge, 1961). The method therefore errs on the conservative side as far as individual stimuli are concerned. Summation of within subject scores, on the other hand, permits one to use parametric techniques.

The scores of the Bukusu and Bualé groups were compared using Student's t-test. They were found not to differ significantly ($t = 0{\cdot}406$; d.f. = 37; $P > 0{\cdot}10$) and were therefore combined for the purpose of further analysis. The combined score was compared with the zero value which would occur in the absence of implicit constancy. By the same test the difference was found to be significant ($t = 2{\cdot}946$; d.f. = 38; $P < 0{\cdot}005$).

It appears therefore, prima facie, that the phenomenon is present in the subjects tested. It could be argued, however, that the observed result is due not to the constancy effect fostered by the presence of the 'cube' but rather to the 'cube's' very presence making it more difficult for the subjects to recognize the correct pattern.

To check on the above possibility a control group of 14 Bukusu subjects was tested using a procedure identical to that described above but with the control set of response choice drawings (Fig. 3).

If the above hypothesis holds these stimuli should evoke various available responses about as often as the original stimuli.

The data obtained were compared with the grouped data obtained in Expt. I. The difference was found to be significant and in the predicted direction ($t = 2{\cdot}696$; d.f. = 41; $P < 0{\cdot}01$).

The evidence therefore suggests that the 'implicit-shape constancy' offers the

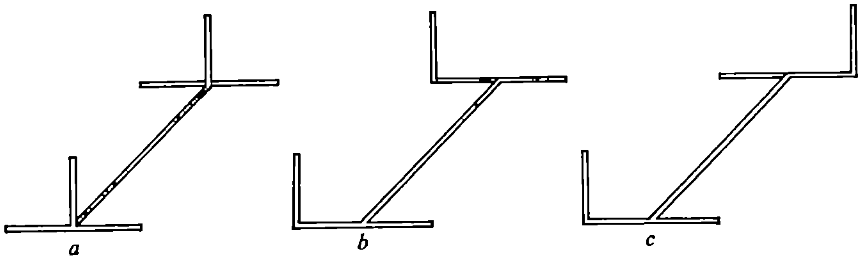


Fig. 4. Drawings used as stimuli in construction task.

most plausible explanation of the obtained results. The alternative explanation which was suggested is rejected.

One would expect those subjects who responded in a manner indicative of implicit perceptual constancy to show other evidence of perception of pictorial depth. To this end the following experiment was carried out.

EXPERIMENT II

Hypothesis

That those subjects who show relatively higher tendency to respond to implicit-shape constancy will also show other evidence of pictorial depth perception more readily than those subjects who exhibit lesser influence of implicit-shape constancy.

Subjects

A random sample of 15 Bualé subjects from Expt. I was used.

Materials

A quantity of Plasticine balls of about 1·5 cm in diameter, ten 0·5 cm in diameter round rods, 15 cm long and sharpened at both ends, and twenty similar rods 7 cm long were used as response materials.

A model of a tetrahedron built with similar Plasticine balls and similar 1·5 cm rods and the three drawings shown in Fig. 4 served as stimuli.

Procedure

The subjects were tested individually. As an introductory task each subject was required to build a tetrahedron similar to that displayed in front of him on the table. The model and the materials were pointed out to him and he was simply told to build a figure like that.

On completion the model was dismembered and one of the drawings was placed in front of him (in a plane inclined at about 30° to the horizontal) and he was requested to build the thing shown in the picture. The response produced was judged by the experimenter to be either two- or three-dimensional. If all the members of the construction lay in the horizontal plane the response was judged to be 'two-dimensional' (2D); otherwise it was judged to be three-dimensional (3D). The essential nature of the judgement criteria was therefore similar to those used by Deręgowski (1968) in his study of pictorial depth perception in Zambian children.

Each subject responded to the three stimulus drawings, which were presented in a random order.

The above task is in general very similar to that used by Deręgowski (1968). It differs from it in one important aspect, namely that of the stimulus figures. It was thought that since Deręgowski's figures could be conceived as vestiges of cubes they would be too closely related to the patterns used in the first experiment to provide means of independent assessment of perception of pictorial depth. It was therefore decided to use a set of alternative figures which, like the figures used in Expt. I, possess but one pictorial depth cue, that of position in the field. If a subject made three 2D responses he was judged to be a 2D-perceiver, otherwise he was judged to be a 3D-perceiver. The entire group was thus divided into nine 3D-perceivers and six 2D-perceivers.

The scores of these two subgroups of subjects obtained in the course of the first experiment were compared and were found to differ significantly in the predicted direction, the 3D-perceivers obtaining scores showing higher tendency towards constancy ($t = 1{\cdot}814$; d.f. $= 13$; $P < 0{\cdot}05$, one-tailed).

Discussion

The results suggest that implicit constancy is not unknown in the populations investigated and that it is correlated with the ability to perceive pictorial depth.

There are unfortunately no data to show how implicit-shape constancy relates to shape constancy proper. However, it does not seem unreasonable to assume that the same mechanism is involved in both phenomena. Granted that this is so, the only data known to the present author and of relevance to the reported findings are two papers by Beveridge (1935, 1939) describing work done in Western Africa. Beveridge found that West African subjects showed greater degree of shape constancy (regression to a circle) than British subjects. Following Thouless's (1933) suggestion educed from observations carried out on Indian subjects he argued that such a tendency was likely to be reflected in the preferred style of art, and postulated that his own data which do not support such a view fail to do so because of the exposure of his subjects to western art styles.

An analogous argument about the relationship between constancy and pictorial perception was subsequently raised by Mundy-Castle & Nelson (1962), albeit in a more precise form. They studied *size* constancy in a remote population. Their argument, however, can probably, *mutatis mutandis*, be applied to *shape* constancy as well. The contention was that groups showing lower Oldfield (1960) constancy scores lack a conceptual bias and made stimulus bound responses to a larger extent than groups having higher values of the index. This, it is claimed, is supported by the scores such groups attain on Hudson's test of pictorial perception (Hudson, 1960), wherein the more stimulus bound subjects fail to show perception of pictorial depth as readily as the less stimulus bound subjects.

In terms of the implicit-shape constancy this argument would run as follows. Subjects who are less influenced by the context in which the stimulus is presented should also find it more difficult to perceive pictorial depth. Our results gratifyingly bear this out.

The results obtained stress the problem of influence which various pictorial elements exert upon each other. There is evidence (Deręgowski *et al.*, 1972) showing that relatively homogeneous figures are recognized, albeit with difficulty, even by subjects with minimal pictorial sophistication. The difficulty which such subjects encounter appears to be that of linking various pictorial elements together to form a cohesive whole, say, a familiar animal. If this difficulty extends, and Hudson's test results suggest that it does, to integration of other depth cues such as, for example, perspective and overlap, which constitute the implying elements of the implicit-shape constancy phenomenon, then one can perhaps regard implicit-shape constancy as an index of pictorial depth perception.

It must be noted that implicit-shape constancy, as defined by the conditions of the present experiment, involves memory storage of the perceived shape. It is therefore not certain that the results obtained here would be observed if the subjects were required to match two stimuli exposed simultaneously. Under those conditions the constancy effect might be easily suppressed by the subjects and play a lesser role than described above. It seems, however, unlikely that it would play no role at all even in those populations which do not tend to make stimulus bound responses. One is inclined to argue that even if the effect were entirely unobservable under conditions of simultaneous presentation where subjects are clearly comparing pictures, it would be merely suppressed but not eradicated.

I am grateful to the S.S.R.C. for a grant for cross-cultural investigation of acquisition of perceptual skills; to Dr Pierre Dasen and Dr Marguerite Lavallée in the Ivory Coast, and to Mr A. Bentley in Kenya, for their help and hospitality.

REFERENCES

BEVERIDGE, W. M. (1935). Racial differences in phenomenal regression. *Br. J. Psychol.* **26**, 59–62.

BEVERIDGE, W. M. (1939). Some racial differences in perception. *Br. J. Psychol.* **30**, 57–64.

BLAKEMORE, C. (1973). The baffled brain. In R. L. Gregory & E. H. Gombrich (eds.), *Illusion in Nature and Art*. London: Duckworth.

DEREGOWSKI, J. B. (1968). Difficulties in pictorial depth perception in Africa. *Br. J. Psychol.* **59**, 195–204.

DEREGOWSKI, J. B. (1971). Pictorial perception and culture. *Scient. Am.* **227**, 82–88.

DEREGOWSKI, J. B. (1973). Illusion and culture. In R. L. Gregory & E. H. Gombrich (eds.), *Illusion in Nature and Art*. London: Duckworth.

DEREGOWSKI, J. B., MULDROW, E. S. & MULDROW, W. F. (1972). Pictorial recognition in a remote Ethiopian population. *Perception* **1**, 417–425.

EKMAN, G. & JUNGE, K. (1961). Psychophysical relationships in visual perception of length, area and volume. *Scand. J. Psychol.* **2**, 1–10.

GREGORY, R. L. (1966). *Eye and Brain*. London: Weidenfeld & Nicolson.

GUERRY, V. (1970). *La Vie quotidienne dans un village Baoulé*. Abidjan: L'Institut Africain pour le développement économique et social.

HUDSON, W. (1960). Pictorial depth perception in subcultural groups in Africa. *J. soc. Psychol.* **52**, 183–208.

HUDSON, W. (1962). Pictorial perception and educational adaptation in Africa. *Psychol. Afr.* **9**, 226–239.

HUDSON, W. (1967). The study of the problem of pictorial depth perception. *Int. J. Psychol.* **2**, 89–107.

MILLER, R. J. (1973). Cross-cultural research in the perception of pictorial materials. *Psychol. Bull.* **80**, 135–150.

MUNDY-CASTLE, A. C. & NELSON, G. K. (1962). A neuropsychological study of the Knysna forest workers. *Psychol. Afr.* **9**, 240–272.

OLDFIELD, R. C. (1960). A simplified method of treatment of size constancy data. *Q. Jl exp. Psychol.* **12**, 33–35.

THOULESS, R. H. (1933). A racial difference in perception. *J. soc. Psychol.* **4**, 330–339.

(*Manuscript received* 11 *November* 1974; *revised manuscript received* 17 *April* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 31–45
Printed in Great Britain

SOME EVIDENCE FOR 'SPEECH' AS AN ACOUSTIC FEATURE

By JOHN MORTON AND SUSAN M. CHAMBERS
Medical Research Council Applied Psychology Unit, Cambridge

Under conditions of serial recall of auditorily presented lists of digits, recall of the last item has been shown to be adversely affected by the presence of a redundant item following the list. This is known as 'the suffix effect' (Crowder & Morton, 1969). In a series of experiments it is shown that the size of this effect is not influenced by the phonological complexity of the suffix. Non-speech sounds, on the other hand, produce no suffix effect even when the subjects are forced to process them. Certain speech sounds were also found to produce no effect. It is concluded that these sounds lacked properties which are characteristic of speech sounds and so were classified as 'non-speech' and that as a result, these sounds are processed by a separate system from the speech sounds.

It has long been known that when a supra-span list of unconnected items is presented for serial recall (i.e. recall in the order of presentation) there is an advantage for auditory over visual presentation. This advantage is restricted to the last few items in the list. Conrad & Hull (1968) and Routh (1970) have shown a similar advantage for visually presented lists when the subjects rehearse each item aloud at the time of presentation. Murray (1965) showed that the advantage given by vocal rehearsal is removed when the subject's voice is masked from himself by means of white noise. Thus it would appear that we are talking about a specifically acoustic effect and not something to do with such factors as ease or force of rehearsal. Crowder (1967) and Morton (1968) have also shown that the advantage of auditory over visual presentation can be removed by a *stimulus suffix*, an extra irrelevant or redundant item presented immediately after the last item in the stimulus list. We have, then, two characteristic shapes of error function which are shown in Fig. 1. With normal acoustic presentation or with visual presentation with vocal rehearsal we have the 'acoustic' curve in which the number of errors on the final item is approximately equal to the number of errors made on the initial item. With visual presentation, or with acoustic presentation with a stimulus suffix we have the 'visual' curve in which the number of errors on the final item is roughly equal to the errors made on items in the middle of the list.

On the basis of these results Crowder & Morton (1969) have postulated the existence of a precategorical acoustic store (PAS). This is conceived of as being a property of that part of the nervous system responsible for the extraction of phonological features from a speech input. There is a certain amount of evidence concerning the location of PAS (using the term 'location' functionally not anatomically). This evidence has been obtained from experiments involving the use of a stimulus suffix (hereafter simply 'suffix'). Thus Morton *et al.* (1971) have shown that if the stimulus is presented in one ear, then a suffix in the same ear (ipsilateral) has more effect on recall than a suffix in the opposite ear (contralateral). As a binaural suffix has the same effect as a contralateral suffix it is clear that the suffix is not ear-specific but is rather channel-specific in the sense of Broadbent (1958). Further

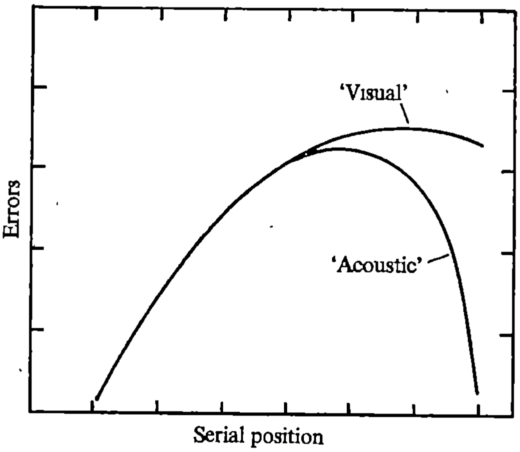


Fig. 1. Characteristic shapes of error functions with serial position for serial recall of lists of auditorily and visually presented items.

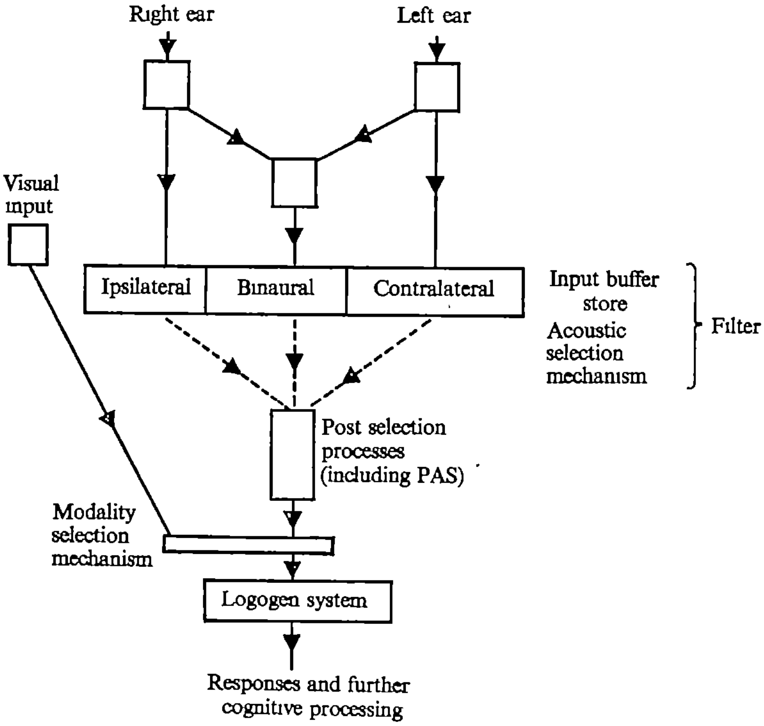

Right ear
Left ear
Visual input
Ipsilateral
Binaural
Contralateral
Input buffer store
Acoustic selection mechanism
Filter
Post selection processes (including PAS)
Modality selection mechanism
Logogen system
Responses and further cognitive processing

Fig. 2. Information-processing model of precategorical processes involved in the suffix effect.

experiments by the same authors indicate that PAS must largely be a property of processes which follow those mechanisms implicated in the selection of a particular acoustic channel. Furthermore, since Crowder & Morton (1969) and Morton & Holloway (1970) have shown that an acoustic suffix has no selective effect on recall of the final items of a visually presented list, and a visual suffix has no effect on recall of an acoustically presented list, even when the full processing of the suffix was guaranteed, then we can conclude that PAS must be located before the convergence of auditory and visual information. These relationships are summarized in the flow diagram given in Fig. 2. This diagram has as a context a model for word recognition which was derived independently of considerations of memory (Morton, 1969). One implication of the model is that the semantic properties of the suffix should have no influence on the size of the suffix effect. This prediction has been confirmed by Morton *et al.* (1971).

Another acoustic variable which affects the action of the suffix is the relationship between the voices used for suffix and stimuli. If one is a male voice and the other a female voice the size of the suffix effect is reduced (Morton *et al.*, 1971, expts. XIV, XV). The suffix paradigm is clearly responsive to acoustic features of the suffix. Having the suffix in a different spatial channel or voice from the stimuli merely reduces the suffix effect. If a burst of white noise is used as a suffix, however, the suffix effect vanishes completely (Morton *et al.*, 1971, expt. XIII). There are a number of ways in which a burst of noise differs from the usual suffix – a spoken 'nought' or 'zero', and the purpose of the current series of experiments was to determine precisely the nature of the crucial variable. Two clear candidates were the phonetic complexity of the suffix and the extent to which the suffix resembled natural speech.

General method

Groups of subjects listened to lists of seven digits. The digits were recorded at a rate of 2 per sec. on a Vortexion tape-recorder and played through the recorder's loudspeaker. In all the experiments the basic method was to present lists of digits for serial recall under a number of conditions – the conditions corresponding to suffixes of different types. Unless otherwise mentioned, the conditions were randomized in a block of 54 stimulus lists, each condition being represented by nine stimulus lists within which all digits occurred equally often at all serial positions. In those conditions involving a suffix, the suffix followed the final digit at an interval of 0·5 sec. The test items were preceded by 16 practice lists which illustrated all the conditions which were to follow. Subjects were instructed to recall the digits by writing them down in their order of occurrence on prepared response sheets. They were requested not to leave any blanks and had 12 sec. to make their responses before a 'ready' signal for the next list. Each experiment took about 30 min. Only those responses in the correct serial position were scored correct. The subjects, all members of the APU subject panel, were females aged between 21 and 67. Subjects who made no errors at all under any one condition were excluded from the data analysis and do not appear in the subject numbers given below. This precaution was necessary as the pool of subjects had widely differing abilities, and for those subjects whose span exceeds eight items there could be no suffix effect under any conditions. It should be noted that the requirement for the subjects to recall the items in order of presentation (serial recall) is essential when the effects of suffixes are being investigated. In a short-term memory paradigm such as the one being used there appear to be a number of sources of information which the subject can use. While Crowder & Morton (1969) and Morton *et al.* (1971, expt. I) have demonstrated the effects of a suffix with

Table 1

Harmonic ...	1	2	3	4	5	6	7	8
Natural	17	16	23	24	31	26	0	10
Synthetic	20	22	28	31	34	31	26	7

other methods of recall, the effects are much smaller than those found with serial recall. Thus subjects were carefully watched to ensure the appropriate recall order was used.

EXPERIMENT I

There were six conditions involving variations in the suffix. It was expected that the different suffixes would give different effects as a function of their similarity to speech. In all cases the signal to be used was recorded on a tape loop. The stimulus items had been recorded on one channel of a two-channel Vortexion tape recorder in a female voice. This channel was played and a timing and gating circuit was arranged such that a 310 msec. segment of the appropriate suffix tape loop was recorded on the second channel 0·5 sec. after the onset of the last stimulus item. The rise and fall times of the gate were about 2 msec. The six suffixes used were:

1. *Natural 'AH' I.* A natural steady-state vowel /a/. of the same fundamental frequency as the stimulus items (210 Hz).
2. *Synthetic 'AH'.* A synthetic steady-state vowel sound made by passing a pulse source of 210 Hz through two formant filters adjusted to give approximately the same vowel quality as the natural vowel in condition 1. These two sounds were analysed by a Kay Sonograph and their harmonic components are given in Table 1.
3. *White noise I.* White noise was passed through the formant filters as set for condition 2. This produces a sound like a whispered /a/.
4. *Buzz.* The pulse train used in 2.
5. *White noise II.* Unfiltered white noise.
6. *Tone.* A pure tone of 210 Hz.

Once the two-channel recording had been prepared, five judges listened to the tape over a loudspeaker with the channels mixed and the playback levels of the different suffixes were adjusted until they appeared to be of equal loudness to the stimulus items. The stimuli and suffixes were then re-recorded on to a single channel maintaining the appropriate relative levels.

Eleven subjects listened to the recording in a single group. They were told 'You will hear lists of digits, each list followed by a noise. When you hear the noise you are to begin your recall.' The rest of the experimental method is described above under 'General Method'.

A further group of ten subjects listened to the same tape with slightly different instructions. They were told, 'After the digits you will hear one of a number of sounds – one of them me saying "ah".' We wished to ensure that the explicit reference to speech would make no difference to performance.

Results

There were no differences in performance attributable to the instructions. The differences between the conditions for the two groups were tested with the Wilcoxon test (Siegel, 1956, p. 75) at all serial positions. With 15 comparisons and seven posi-

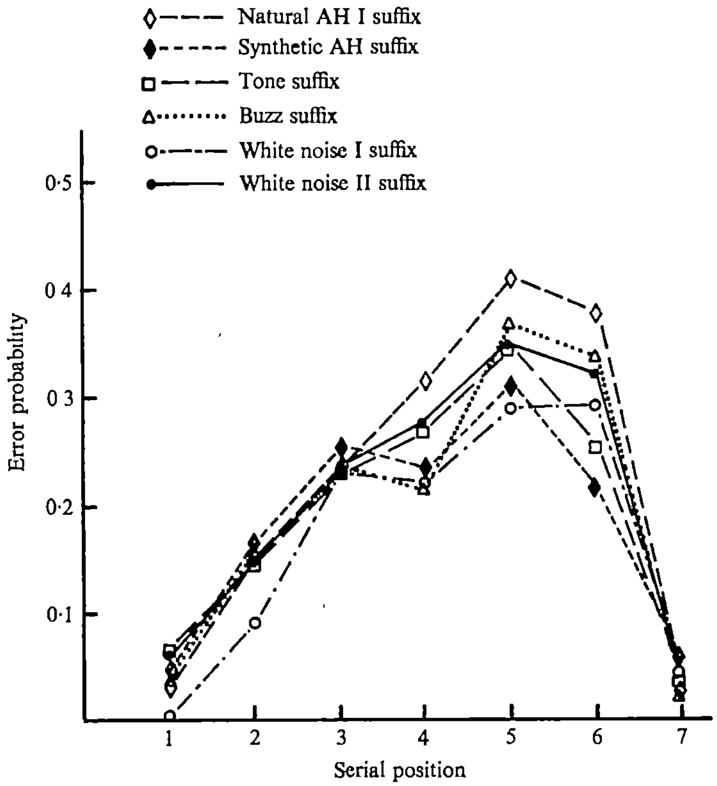


Fig. 3. Error probabilities by serial position for a variety of non-speech suffixes and a gated natural 'ah'. (Expt. I, $n = 21$.)

tions for each there were a total of 105 comparisons of which one was significant at the 1 per cent level, one at the 2 per cent level and five at the 5 per cent level. As the differences were completely unsystematic they were regarded as being the results of random variations. In addition there were no suffix effects in any of the conditions, the recall curves being typically 'acoustic' in shape. The mean error curves for the two groups are shown in Fig. 3.

Discussion

This result was rather surprising as we had every reason to suppose that a spoken vowel sound would produce a normal suffix effect. We supposed then that our expectations were incorrect and that a single vowel sound was too simple to produce a suffix effect. The next experiment was designed to include suffixes of increasing phonological complexity.

EXPERIMENT II

The stimulus lists were those used in the preceding experiment. The suffixes were spoken onto one channel of a tape recorder as the stimulus lists were being transcribed onto the other channel. The two channels were later mixed with the suffixes adjusted to the loudness level of the stimuli. The six conditions were:

1. 'Ah' – /a/.
2. 'Oh' – /ou/.

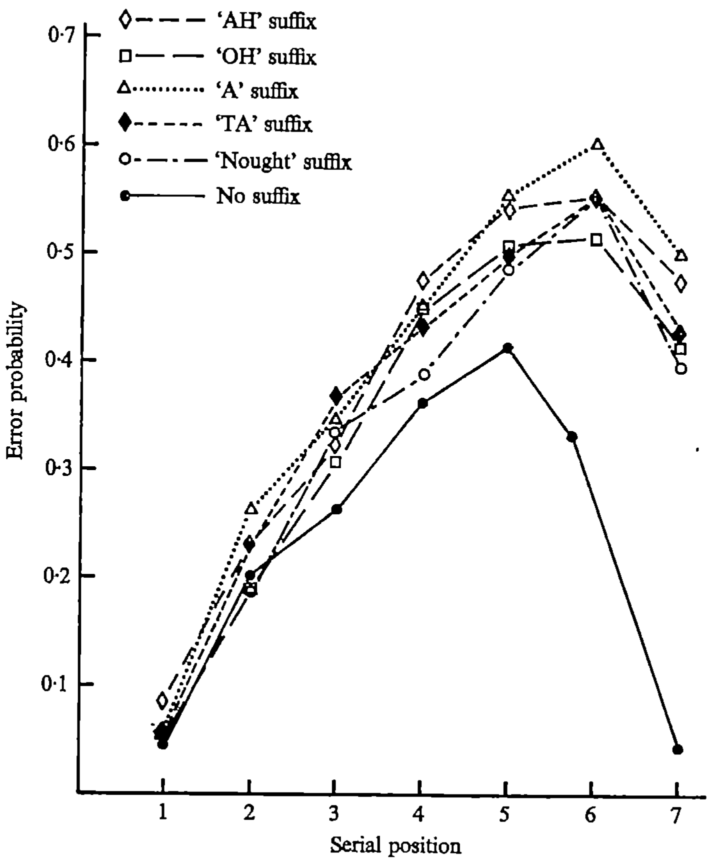


Fig. 4. Error probabilities for a variety of speech suffixes. (Expt. II, $n = 29$.)

3. 'A' – /eI/.
4. 'Ta' – /ta/.
5. 'Nought' – /not/.
6. Control (no suffix).

Previous studies had shown that 'nought' gave a large suffix effect and we expected the other suffixes to give results between that and the control condition. Eleven subjects listened to the stimuli in a single group. They were told 'you will hear a spoken sound following each list of digits. Occasionally there will be nothing following. Begin recall as soon as the sound has occurred or as soon as the sound hasn't occurred.' Any confusion arising from these instructions was dispelled during the practice items.

A further 20 subjects listened to a similar tape on which the pairing of suffixes and lists was different.

Results

As there were no differences between the groups, data from all subjects was pooled. Wilcoxon tests showed that the only significant differences on the last serial position were between the control curve and all other conditions ($P < 0{\cdot}01$, two-tailed). There were no significant differences among the suffix conditions on other serial positions. The pooled error curves are shown in Fig. 4.

The differences between the control curve and the suffix conditions at positions other than the final one are attributed to the suffix having an additional effect resembling that of a response prefix. This effect is considered to be unrelated to the suffix effect (see Morton *et al.*, 1971).

Our predictions in this experiment were that the more complex suffix would have a greater effect. Statistically there were no differences among the suffixes and, in fact, the simplest suffix 'ah' gave the largest effect and the most complex, 'nought' the smallest, though we are inclined to ascribe such differences as exist totally to chance. Since we are aware of no procedural artifact that could disguise a real effect we are willing to conclude that phonological complexity *per se* has no influence on the size of the suffix effect. (Note that in more recent unpublished experiments we have found that phonological similarity between the suffix and the final item in the list increases the size of the suffix effect. Such a phenomenon may interact with phonological complexity.)

Discussion

The results of Expt. II demonstrated that the phonological complexity of the suffix had no systematic influence on the size of the suffix effect. In Expt. I we found that the similarity of the suffix to a speech sound was immaterial; none of the sounds produced a suffix effect. The results of these two experiments appear to be in conflict. In Expt. I the suffix 'ah' had no effect; in Expt. II it had a great effect. The most obvious difference between the sounds is that the one used in Expt. I was initially recorded on a tape loop and then gated on to the stimulus tape with rise and fall times of only about 2 msec. With a naturally spoken vowel the rise times are more typically of the order of 40 msec. and the fall times are even more extended. Tracings of the envelopes of examples of the two sounds are shown in Figs. 5 (1) and 5 (2) respectively. A further experiment, in which naturally spoken 'ah' suffixes were truncated, revealed that truncation *per se* had no effect on the size of the suffix effect. In this experiment, naturally spoken vowel suffixes were gated at the beginning, at the end or both. The cut-off points chosen are illustrated in Fig. 5 by the marks *f* and *b*. The three kinds of suffix which resulted from the procedure were indistinguishable in their effects from an untreated vowel. A number of other experiments established that there were no artifacts in the earlier experiments as a result of instructions or the general 'context' of the suffix sounds (i.e. speech sounds *v.* non-speech sounds).

The source of the conflict was finally settled in the following experiment in which, for a number of now irrelevant reasons, we presented subjects with the same tape that was used in Expt. I but with a new gated 'ah' in place of one of the noise conditions.

EXPERIMENT III

We first made another loop 'ah' in the usual way. Then we inserted a gated portion of this loop in place of the White Noise I suffix in the tape we had used for Expt. I. The resulting tape was played to a group of 22 subjects, with the instructions 'The digits are followed by a noise, sometimes me saying "ah" ', and the usual procedure.

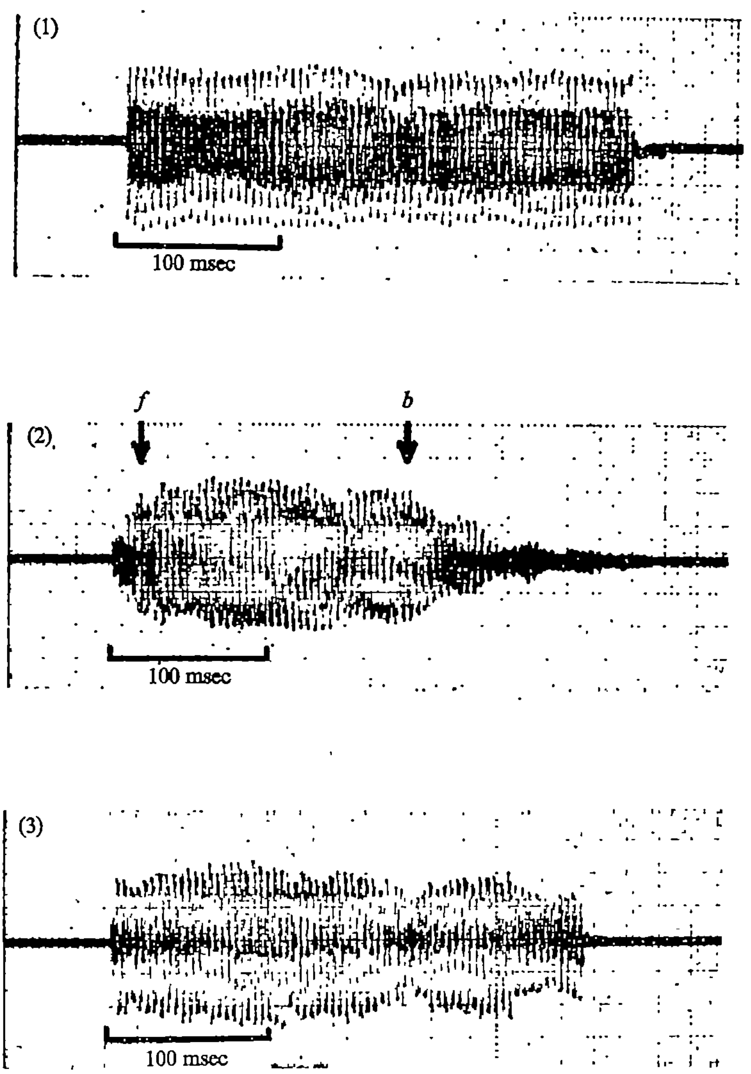


Fig. 5. Amplitude wave-forms for some of the suffixes used. (1) A gated portion of an extended spoken 'ah'. This sound failed to produce a suffix effect in Expts. I and III. (2) A natural spoken 'ah' showing the usual rise and fall characteristics. The arrows labelled *f* and *b* refer to truncation points typical of those used in the experiment described in the discussion of Expt. II. (3) A gated portion of an extended spoken 'ah'. This sound produced a suffix effect in Expts. III and IV.

Results

The mean error probabilities are plotted in Fig. 6. In this figure the new loop is termed 'Natural AH II'. Wilcoxon tests showed differences at the 1 per cent level between this condition and all other conditions on the last two serial positions. Performance in the condition using the original loop 'ah', termed 'Natural AH I' in the figure, is not significantly different from the remaining conditions even at the 5 per cent two-tailed level. Thus it appears that there are significant differences between the two natural loop suffixes.

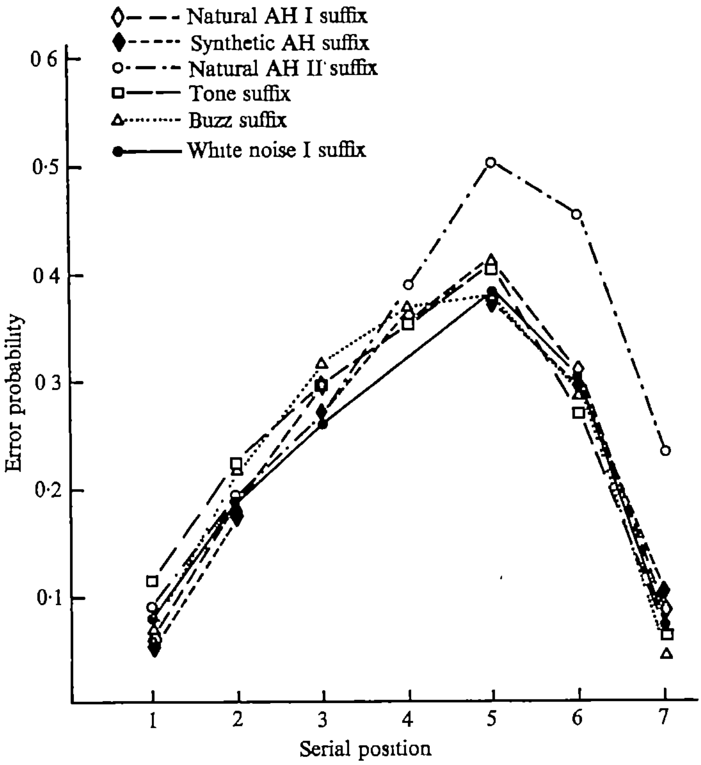


Fig. 6. Error probabilities for a variety of suffixes including one gated spoken 'ah' which produces a suffix effect and one which does not (Expt. III, $n = 22$).

EXPERIMENT IV

Seeking to convince ourselves that there were no remaining procedural artifacts we made two more changes to the original recording, removing the white noise suffix to produce a control condition and replacing the original loop 'ah' by a normally spoken 'ah'. Our reason for running this experiment was that it was possible that the instructions used in Expt. III led the subjects to pick out one particular suffix (on the basis of some as yet unknown cues) and process this differently. This selection procedure could happen during the practice lists and could have been fully established by the beginning of the test lists. If this is the reason for the suffix effect found with Natural AH II in Expt. III, then in this experiment we would expect the normally spoken 'ah' to be selected and then would expect to get no suffix effect in the condition involving Natural AH II. We had used the same instructions with the second group in Experiment I and had found no effect with the original Natural 'ah' loop, but it was possible that other, unidentified procedural variations had crept in to the running of the experiments. The control condition was included to make sure that we had an unambiguous baseline against which to compare the effects of all the suffixes. A group of 20 subjects were used in the experiment which in instructions and procedure was identical to Expt. III.

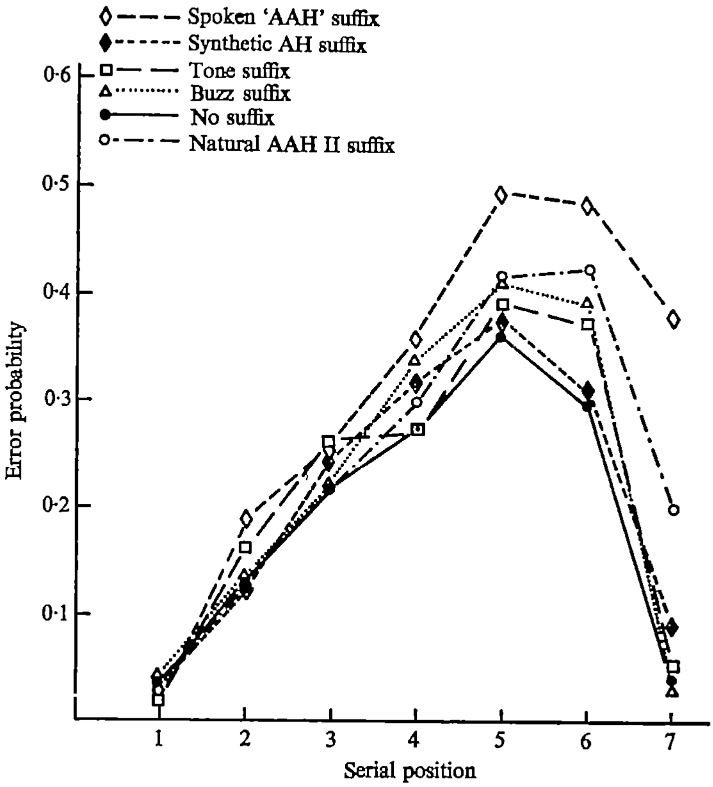


Fig. 7. Error probabilities for a variety of suffixes including a spoken vowel and a gated vowel (Expt. IV, $n = 20$).

Results

The mean error probabilities are plotted in Fig. 7. Wilcoxon tests showed significant differences on the final serial position between the spoken suffix and all other conditions and between the natural (loop) suffix and the remaining conditions. Both the spoken and the natural suffixes gave rise to more errors than the control condition on serial position 6. This experiment confirms that the difference between the gated natural vowels which was found in Expt. III was not a consequence of the particular instructions used. One consistent feature of the last two experiments is that the loop suffix used had about the same effect, giving rise to 24 per cent errors on the last position in Expt. III and 20 per cent in Expt. IV. In this it contrasts not only with the original loop suffix, which gave rise to only 2 per cent errors in Expt. I, but with the figure of 40–50 per cent errors found in Expt. II with normally spoken suffixes. These differences are unlikely to be due to differences in ability between the subjects in the various experiments since the errors made in the first four serial positions follow the same pattern in all the experiments. We are thus left only with the possibility that some vowel sounds which are sustained for about 10 sec. may have characteristics of synthetic sounds.

One way of describing the difference between the sounds is in terms of their 'naturalness'. The synthetic vowel sound we used in Expt. I was unambiguous in its quality, being unmistakably an 'ah' sound, but was nonetheless clearly not a human sound. On listening to the natural loop 'ah' we had used in the first experiment it became apparent that it, too, in some way was not quite 'natural' in spite of its human origins. The next experiment was performed to verify our own intuitions.

EXPERIMENT V

The object of the experiment was to compare the 'naturalness' of the spoken, loop and synthetic suffixes used in Expts. I and II. There were nine samples of the loop suffix and 18 of the other two. These 45 suffix sounds were re-recorded from the original tapes on to another tape in a semi-random order at 6 sec. intervals. Each sound was recorded twice, giving a total of 90 test stimuli. A group of 13 subjects were asked 'to give each sound a rating as to how natural it seems'. They were informed that the experiment had to do with the assessment of different recording and reproduction equipment. The responses were made on prepared sheets which had a seven-point scale to be marked for each stimulus. On this scale 'Natural' corresponded to a rating of 1 and 'Unnatural' to a rating of 7. Prior to the test items there was a block of 15 items, taken from the practice blocks of the same experiments. Unknown to the subjects, these items were treated as practice items.

Results

The ratings were treated in two ways. The distribution of the modal ratings for each stimulus item is shown in Table 2. It is clear from this table that the subjects considered the loop suffix to be closer to the synthetic than to the spoken 'ah'. The modal response of each subject to the three different sounds was also obtained. All 13 subjects had a lower rating of the spoken sound than the other two, and all except one subject rated the loop sound more natural than the synthetic. The other subject had the same *modal* rating to both sounds but a lower *mean* rating to the loop sound.

EXPERIMENT VI

The previous experiment showed that the loop 'ah' which did not produce a suffix effect was in some way 'unnatural'. We decided to investigate the 'naturalness' rating method further to see if it was more sensitive than the suffix effect. In the discussion to Expt. II we mentioned an experiment in which no differences had been found between the suffix effects of vowel sounds truncated in a number of ways. A tape was prepared with all the sounds used in that experiment. There were nine examples of each of five types of suffix. They were arranged in random order on the tape and played to a group of 11 subjects under conditions similar to those used in Expt. V.

Results

The mean ratings for the five sounds are given in Table 3. Differences between the sounds were tested by obtaining a mean rating for each group of sounds for each subject. Differences between these values were then tested by the Wilcoxon test. The significance levels are also given in Table 3. It is clear from this table that the rating method distinguishes between sounds which do not differ in their effectiveness

Table 2

	Distribution of modal ratings for individual subjects							Mean ratings	Distribution of modal ratings for individual items						
	1	2	3	4	5	6	7		1	2	3	4	5	6	7
Spoken	3	10	—	—	—	—	—	1·76	14	21	1	—	—	—	—
Loop	—	—	2	2	7	2	—	4·60	—	—	1	5	8	4	—
Simulated	—	—	—	2	5	8	2	5·81	—	—	—	—	3	26	7

Table 3

		Significance of differences with other ratings			
Suffix	Mean rating	Back	Front	Front–Back	Loop
Natural	3·08	n.s.	0·05	0·02	0·01
Back truncated	3·10	—	0·01	0·01	0·01
Front truncated	3·41	—	—	n.s.	0·02
Front and back truncated	3·80	—	—	—	n.s.
Loop	4·32	—	—	—	—

as suffixes. This might be a result of a greater sensitivity in measuring the same variable. Alternatively it might be that the characteristics of sounds which cause them to have a suffix effect and those which lead listeners to rate a sound as more natural than another have only a partial overlap.

Discussion

The following facts appear to have been established.

1. Non-speech sounds do not produce a suffix effect irrespective of instructions concerning their nature or the set of other sounds which could also occur as suffixes (Expts. I, III and IV).

2. The size of the suffix effect with normally spoken sounds is not affected by the phonological complexity of the suffix, the simple vowel sound 'ah' producing as big an effect as the word 'nought' (Expt. II).

3. The naturally spoken vowel sound 'ah' produces a suffix effect irrespective of instructions concerning its nature or the set of other sounds which could also occur as suffixes (Expts. II and IV).

4. When a portion of an extended 'ah' vowel is used as a suffix it may or may not produce a suffix effect. One particular such sound failed to produce a suffix effect under a variety of instructions and contexts (Expts. I, III and IV). Other such sounds did produce a suffix effect (Expts. III and IV).

5. A good quality synthetic vowel produced no suffix effect.

6. When asked to rate for naturalness the loop 'ah' which did not give a suffix effect subjects placed it nearer a synthetic 'ah' than a spoken 'ah' (Expt. V).

7. The effect of a naturally spoken 'ah' is not diminished when the natural rise and fall characteristics are removed.

8. Subjects gave different 'naturalness' ratings to various truncated vowel sounds which had not given rise to differences in suffix effect (Expt. VI). It might be that the naturalness rating is more sensitive than the suffix effect or that the two do not depend on exactly the same parameters.

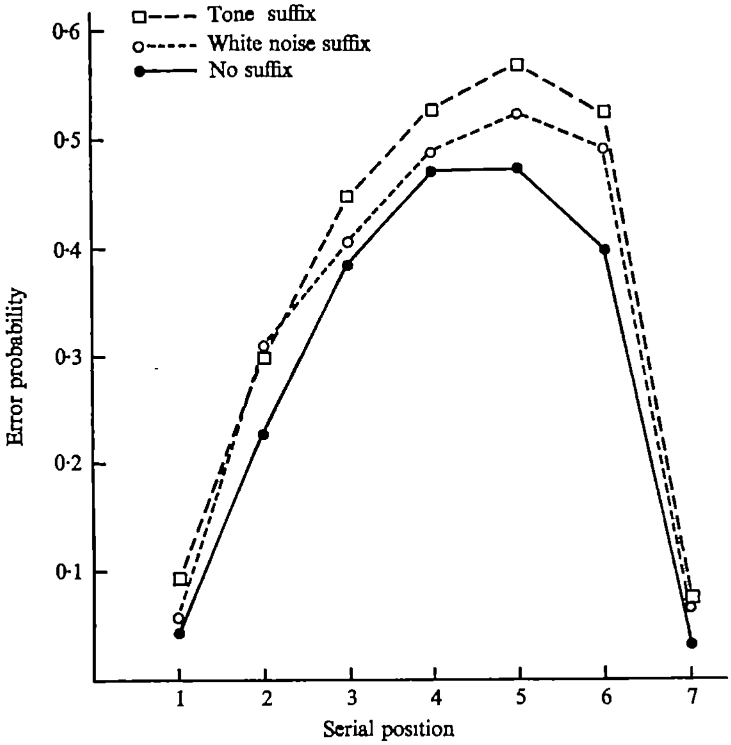


Fig. 8. Error probabilities in a suffix-prefix experiment where subjects had to identify the suffix by writing I for tone, *N* for noise or a cross for no suffix. This procedure failed to induce a suffix effect (Expt. VII, $n = 22$).

The precise nature of the acoustic feature which distinguishes a sound as 'natural' or 'speechlike' and thus gives rise to a suffix effect is unknown. The two main possibilities are the amount of variability in the fundamental frequency of the sound and in the formant frequencies. The natural loop suffix which failed to produce an effect showed very little variability in the waveform with time. This can be seen in Fig. 5 which shows the amplitude waveforms of this suffix, a normal 'ah' and the natural Loop II which did produce a suffix effect. The irregularities in the latter two are clear.

We suppose, then, that incoming sounds are automatically analysed in various ways to establish whether or not they are speech, one of the characteristics of speech sounds being some kind of irregularity. Speech sounds would then be automatically selected and passed through PAS to produce a suffix effect. Non-speech sounds would be routed through another system or completely blocked.

We must note here that it is not possible to suggest that non-speech sounds are processed in PAS but do not interfere with the trace of the final item. Such a statement would be a mixed metaphor. In a functional model non-interference of the kind demonstrated is a sufficient condition for requiring different systems.

When we claim that non-speech sounds do not produce a suffix effect we are, of course, only generalizing from the rather narrow range of sounds we employed. It remains possible that sounds like a cough, a dog barking or the chime of Big Ben could produce an identical effect. We must leave such discoveries to others. Equally, the possibility of a store for non-speech sounds with properties similar to those of PAS remains to be investigated. Clearly non-speech sounds can be stored. Massaro (1970) provides some evidence for storage of pure tones but he obtained contralateral interference effects as great as ipsilateral. PAS cannot be implicated since Morton *et al.* (1971) showed that contralateral effects are greatly reduced. We presume that the brief auditory store described by Treisman & Rostron (1972) on the basis of experiments using pure tones is to be equated with that required by Massaro.

There remains to test the possibility that non-speech sounds are blocked from PAS. It is legitimate to ask whether the non-speech sounds would produce a suffix effect if they have to be processed. The following experiment tests this.

EXPERIMENT VII

Using the same procedure as outlined in the General Method section a group of 22 subjects listened to the digit lists. There were three suffix conditions. In one of them the list was followed by a 300 msec. burst of white noise. In the second condition was a 300 msec. pure tone of 210 Hz. For the third condition there was no suffix.

The main variation from the General Method was that prior to recall the subjects had to indicate the nature of the suffix. This they did by writing *N*, if the suffix was noise, I if it was a pure tone and putting a cross if there was no suffix. Otherwise the conditions of presentation, instructions and scoring was as before.

Results

The mean error curves are shown in Fig. 8. Differences between the conditions were tested by means of the Wilcoxon test. The noise condition differed from the control condition at position 6, and the tone differed from the control at positions 1, 5 and 6 ($P < 0{\cdot}01$, two-tailed). There were no differences between conditions on the final position even at the 5 per cent level.

DISCUSSION

Under conditions similar to those used in this experiment, Morton *et al.* (1971, expt. XII) found that the effect of a suffix presented in the opposite ear to the stimulus was greatly increased compared with the results when the subjects did not have to respond to the suffix. They took this result as evidence that PAS is located after the selection mechanism which operates to reduce the effect of the unattended contralateral suffix. Where the suffix has to be processed the selection mechanism cannot operate and the suffix then affects the information in PAS.

The present experiment shows that non-speech sounds do not produce a suffix effect even when they have to be recognized and responded to. We therefore conclude that PAS is a property only of the speech analysis system and that non-speech sounds, even when fully processed are not processed by the part of the system implicated in the suffix effect.

We are grateful to Byron Morgan for providing a program for the Wilcoxon tests and to Keith Tayler for his kind assistance with the data processing.

REFERENCES

BROADBENT, D. E. (1958). *Perception and Communication*. New York: Pergamon Press.

CONRAD, R. & HULL, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychon. Sci.* **10**, 135–136.

CROWDER, R. G. (1967). Prefix effects in immediate memory. *Can. J. Psychol.* **21**, 450–461.

CROWDER, R. G. & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Percept. Psychophys.* **5**, 365–373.

MASSARO, D. W. (1970). Perceptual auditory images. *J. exp. Psychol.* **85**, 411–417.

MORTON, J. (1968). Selective interference in immediate recall. *Psychon. Sci.* **12**, 75–76.

MORTON, J. (1969). The interaction of information in word recognition. *Psychol. Rev.* **76**, 165–178.

MORTON, J., CROWDER, R. G. & PRUSSIN, H. A. (1971). Experiments with the stimulus suffix effect. *J. exp. Psychol. Monogr.* **91**, 169–190.

MORTON, J. & HOLLOWAY, C. M. (1970). Absence of cross-modal 'suffix effects' in short-term memory. *Q. Jl exp. Psychol.* **22**, 167–176.

MURRAY, D. J. (1965). The effect of white noise on the recall of vocalized lists. *Can. J. Psychol.* **19**, 333–345.

ROUTH, D. A. (1970). 'Trace strength', modality and the serial position effect in immediate memory. *Psychon. Sci.* **18**, 355–357.

SIEGEL, S. (1956). *Non-parametric Statistics*. New York: McGraw-Hill.

TREISMAN, M. & ROSTRON, A. B. (1972). Brief auditory storage: a modification of Sperling's paradigm applied to audition. *Acta psychol.* **36**, 161–170.

(*Manuscript received* 1 *May* 1974;
revised manuscript received 4 *February* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 47–52
Printed in Great Britain

EXTRAVERSION AND REINFORCEMENT IN VERBAL OPERANT CONDITIONING

BY B. S. GUPTA*
Department of Psychology, Guru Nanak University,
Amritsar, India

The present investigation was designed to study the relationship between extraversion and modes of reinforcement in verbal operant conditioning. A 4 × 2 randomized block design was replicated ten times. Eighty graduate and post-graduate male students were individually subjected to the standard Taffel conditioning procedure. When the numbers of conditioned responses produced by introverts and extraverts were compared, it was found that: (1) under negative reinforcement ('bad' and electric shock), with both the female and male experimenters, and under positive reinforcement ('good') with the male experimenter, the introverts' score was the higher of the two; (2) under positive reinforcement ('good') with the female experimenter, the extraverts' score was higher than that of the introverts.

Eysenck's (1947, 1960, 1967*a*) theory of personality involves two major dimensions: extraversion and neuroticism. Extraversion is based on the excitation/inhibition balance and neuroticism is equated with general emotionality. The theory suggests that introverts invariably condition more rapidly than extraverts. The studies of Eysenck (1959*a*), Quay & Hunt (1965), Gupta (1970, 1973), Gupta & Singh (1971), and Hekmat (1971) support this assumption for verbal conditioning. However, on the basis of a study conducted in the Maudsley Laboratory on eye blink conditioning Eysenck & Levey (1972) conclude that introverts display greater conditionability than extraverts only in certain circumstances and that in others, the reverse is true. Similarly Brody (1972) conjectures that introverts condition relatively slowly in anxiety-producing situations because of the excessive operation of transmarginal inhibition. Gray (1970, 1971) hypothesizes that introverts are more susceptible to punishment and extraverts are more sensitive to reward. Thus, according to Gray, the increasing degree of sensitivity to punishment, as compared with the sensitivity to reward, represents the increasing degree of introversion. Introverts are, therefore, expected to display greater conditionability under punishing conditions while the opposite is expected of extraverts. Gray (1970, 1971) regards emotionality or neuroticism as the degree of sensitivity to both reward and punishment. The increasing degree of sensitivity to reinforcement, therefore, represents the increasing degree of neuroticism. Thus Gray hypothesizes that susceptibility to punishment should grow with increasing neuroticism and decreasing extraversion, but that sensitivity to reward should grow with both neuroticism and extraversion. Eysenck (1957) on the other hand agrees with Spence (1956) and considers emotional arousal as analogous to the Hullian (1943) construct of drive. He also states that this dimension probably does not exert its influence in normal subjects unless the situation is perceived as anxiety-producing (Eysenck, 1967*a*, *b*). A recent study on verbal conditioning, on normal subjects, by Gupta & Singh (1971) reveals that neuroticism exerts little influence. Thus the degree of sensitivity to reinforcement alone (when the susceptibility to reinforcement, i.e. reward or punishment, is held constant) does not seem to play any role in a completely non-threatening situation as in Taffel's (1955) verbal conditioning procedure.

---

* The data were collected when the author was at Government College, Kurukshetra, India. The author is extremely thankful to Dr K. C. Sharma, Principal of the College, for providing necessary facilities for data collection.

The present investigation was designed to examine the relationship, if any, between extraversion and modes of reinforcement (reward or punishment) in verbal operant conditioning (VOC). Following Gray (1970, 1971) it was hypothesized that for introverts negative reinforcement would be more effective while for extraverts, on the other hand, positive reinforcement would be more effective. Another objective of the study was to observe whether the sex of the individual who administers the reinforcement also influences the conditionability of the subjects.

METHOD

Subjects

The subjects were graduate and post-graduate male students of different colleges and university teaching departments. Their age-range was 20–25 years. Ten subjects were assigned randomly to each of the four reinforcement conditions at each level of extraversion. The Hindi translation (Singh, 1966) of Eysenck's (1959) Maudsley Personality Inventory (MPI) was used for determining the level of extraversion. The number of subjects was 80. They were selected out of a sample of 1500 students to whom the MPI was administered. The extraversion scores of the selected subjects were: $E+ = 39+$; $E- = 15-$. The selection of the subjects was done on the basis of Mean ± 1·5 S.D. of the extraversion scores (E Mean $= 26·8$; E S.D. $= 7·9$). The Taffelian (1955) verbal conditioning situation is not anxiety-provoking but in order to control and minimize any possible effect of neuroticism, only those subjects were finally selected who ranged between mean ± 1·0 S.D. on this dimension. The means and S.D.s of 1500 students were used for grouping the subjects.

Experimenters

There were two experimenters: one male (the author himself) and the other female (a graduate student in psychology). At the time of data collection the male and female experimenters were 35 and 21 years of age respectively. Prior to the collection of data the female experimenter was given a thorough training in the conditioning procedure until she became as skilled as the male experimenter in presenting the stimulus material and providing reinforcement to the subject. The data were first collected by the female experimenter and then about six months later by the male experimenter on the same subjects.

Stimulus material

Taffel's (1955) sentence completion technique was used. The stimulus material was 100 white unlined index cards (3 × 5 ins). A neutrally toned verb in the past tense, was typed in block letters in the centre of each card. Five pronouns I, WE, YOU, HE and THEY were inserted in one line in block letters below the verb. The sequence of pronouns on each card differed randomly.

Procedure

The subject's task was to construct for each card a sentence containing the verb and beginning with one of the pronouns. The first 20 cards were used to establish the operant level (initial score). The experimenter simply noted the subject's response whenever he began a sentence either with 'I' or 'WE'. For the next 60 cards, reinforcement was given to the subject whenever he used 'I' or 'WE' at the beginning of a sentence. The modes of reinforcement were as follows: (1) Good; (2) Buzzer; (3) Bad; (4) Mild electric shock to the index finger of the right hand kept in a relaxing manner on a table in front of the subject. The word 'Good' or

Table 1. *Results of analysis of variance*

Experimenter	Source	d.f.	M.S.	*F*	*P*
Male	Extraversion (*E*)	1	2·47	4·41	0·05
	Reinforcement (*R*)	3	3·02	5·40	0·01
	$E \times R$	3	2·35	4·20	0·01
	Within	72	0·56		
Female	Extraversion (*E*)	1	2·75	4·10	0·01
	Reinforcement (*R*)	3	2·69	4·01	0·05
	$E \times R$	3	2·85	4·25	0·01
	Within	72	0·67		

Note. The results are based upon two independent analyses.

'Bad' was spoken by the experimenter in a flat unemotional tone. This was the conditioning stage. For the last 20 cards the experimenter did not reinforce. The frequency of the personal pronouns, i.e. 'I' and 'WE' responses, was used as a quantitative measure of the conditioned response. The order in which the cards were arranged was randomized from subject to subject by shuffling them before the commencement of each conditioning session. Immediately after the experiment each subject, tested individually, completed a questionnaire relating to his insight into the response-reinforcement process. The questionnaire was modelled on that used by Miller (1967). All those subjects who were aware of the connexion between their behaviour and the experimenter's reinforcer were eliminated and replaced by others not so aware. The number of such replacements was 30.

RESULTS

The conditioning score for each subject was derived by subtracting the frequency of his sentences beginning with personal pronouns ('I' and 'WE') in the initial block of 20 cards (operant score) from the frequency of such sentences in the final block of 20 cards (test score). In order to make them comparable all the conditioning scores were brought on a common scale by dividing each conditioning score by the corresponding operant score. The final scores thus obtained were named inflexion ratios.

The mean inflexion ratios for the first two reinforcing conditions, i.e. 'Good' and 'Buzzer', were positive and for the last two conditions, i.e. 'Bad' and 'Electric Shock', were negative for both male and female experimenters. This implies that 'I' and 'WE' responses were more frequent following 'Good' and 'Buzzer', and fewer following 'Bad' and 'Electric Shock'. Thus the former two reinforcing conditions acted as facilitators of 'I' and 'WE' responses and the latter two as suppressors of such responses. In the previous literature (Krasner, 1958; Salzinger, 1959) too, the first two reinforcers have been termed as positive reinforcers and the last two as negative reinforcers. (The positive and negative reinforcers will hereafter be known as rewarding and punishing reinforcers respectively.) The ratios thus obtained were subjected to two-way analysis of variance; the results of which are reported in Table 1.

The mean inflexion ratios both for male and female experimenters are presented diagrammatically in Fig. 1. The figure clearly indicates the relationship between levels of extraversion and modes of reinforcement.

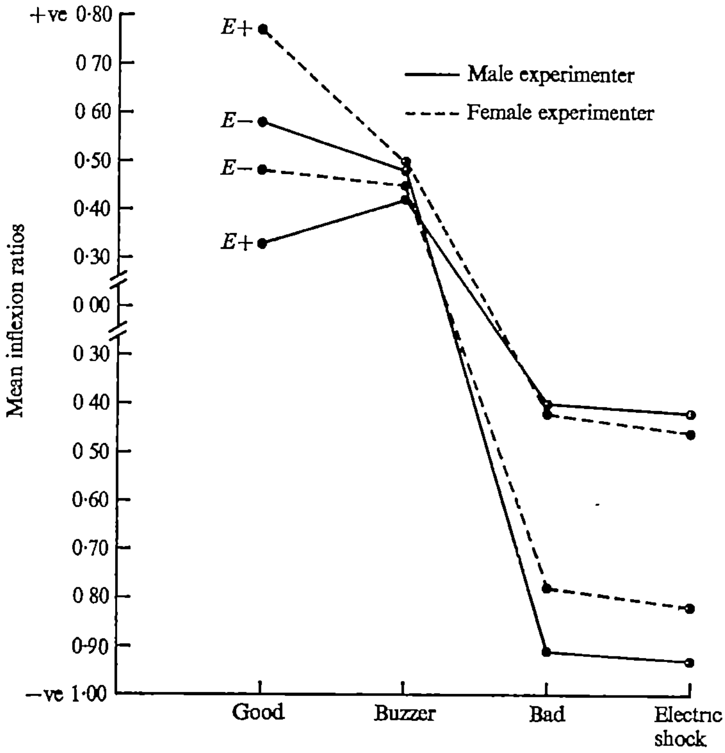


Fig. 1.

DISCUSSION

It is obvious from the results of analysis of variance (Table 1) that extraversion, modes of reinforcement, and their interaction are significant variables for VOC. The significant F-ratio for the interaction of extraversion and reinforcement conditions indicates that the differences between the means of various reinforcement conditions are not independent of the levels of extraversion. Thus Gray's (1970, 1971) assumption that individuals differ in their susceptibility to reinforcement seems valid for VOC. Interaction between the two variables is shown clearly in Fig. 1.

The differences between the means of introverted and extraverted subjects for each of the reinforcement conditions were evaluated by t-test. It was observed that, with the male experimenter under both the punishing conditions ('Bad' and 'Electric Shock'. t-ratios: 'Bad' = 5·54; 'Electric Shock' = 5·26; $P < 0{\cdot}01$ in each case) and under one rewarding condition ('Good'; t-ratio = 3·55; $P < 0{\cdot}01$) the introverts were successfully conditioned more often than were the extraverts. The results for the other rewarding condition ('Buzzer') were in the same direction but were not statistically significant. The results in this section are thus in accord with our earlier observations (Gupta, 1970, 1973; Gupta & Singh, 1971) and support Eysenck's (1947, 1960, 1967*a*) theory to the extent that, in this situation, introverted subjects are more readily conditioned than are extraverts. The results also support Gray's (1970, 1971) assumption that the conditionability of introverts is greater under punishing than under rewarding conditions (*t*-ratios: 'Bad' and 'Good' = 3·67; 'Bad' and 'Buzzer' = 4·68; 'Electric Shock' and 'Good' = 3·76; 'Electric Shock' and 'Buzzer' = 4·50; $P < 0{\cdot}01$ in each case). Similar analysis of mean differences for extraverted subjects was not significant. With the female experimenter (Fig. 1) a different trend was observed. Here the introverted subjects conditioned more often than did the extraverted subjects under punishing conditions (*t*-ratios: 'Bad' = 3·83; 'Electric Shock' = 3·87; $P < 0{\cdot}01$ in each case) but the reverse was true for rewarding conditions ('Good' and 'Buzzer') though the results reached an acceptable level of significance (*t*-ratio = 3·45; $P < 0{\cdot}01$) for one condition ('Good') only. As with the male experimenter, the introverted subjects with the female experimenter conditioned more often under the punishing conditions than under the rewarding conditions (*t*-ratios: 'Bad' and 'Good' = 3·49; 'Bad' and 'Buzzer' = 3·59; 'Electric Shock' and 'Good' = 3·70; 'Electric Shock' and 'Buzzer' = 3·78; $P < 0{\cdot}01$ in each case); the extraverted subjects on the other hand conditioned more often under a rewarding condition ('Good') than under punishing conditions (*t*-ratios; 'Good' and 'Bad' = 3·80; 'Good' and 'Electric Shock' = 3·65; $P < 0{\cdot}01$ in each case). The results for the other rewarding condition, i.e. 'Buzzer' were in the same direction but were insignificant. It seems, therefore, that the word 'Good' spoken in the VOC situation by a young female experimenter is perceived to be more rewarding in nature by equally young male extraverted subjects in comparison to the same word being spoken in the same situation by a male experimenter. Thus Gray's (1970, 1971) assumption that introverts are more susceptible to punishment is supported by our results both with male and female experimenters; his other assumption that extraverts are more sensitive to reward is supported by our results with the female experimenter only. Probably a positive reinforcement which is more rewarding in nature is needed to make the extraverts' condition more often with the male experimenter. The strength of conditioning is to a certain extent determined by an individual's subjective attitude towards the person who administers the reinforcement (Rosenthal, 1966, 1969). At least this is true with our extraverted subjects who show significant differences in conditionability with the male and female experimenter under a rewarding condition ('Good'). Hence both the nature of reinforcement and the experimenter's sex seem to be important for VOC.

It has already been reported that the extraverted subjects conditioned less often than did the introverted subjects under punishing conditions. In other words they made more 'I' and 'WE' responses under punishing conditions than the introverted subjects. This indirectly lends support to the hypothesis of greater pain-tolerance in extraverts than in introverts (Lynn & Eysenck, 1961).

REFERENCES

BRODY, N. (1972). *Personality: Research and Theory*. New York: Academic Press.

EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1959*a*). Personality and verbal conditioning. *Psychol. Rep.* **5**, 570.

EYSENCK, H. J. (1959*b*). *Manual of Maudsley Personality Inventory*. London: University of London Press.

Br. J. Psychol. (1976), 67, 1, *pp.* 53–61
Printed in Great Britain

EXTRAVERSION, ACTIVATION AND THE RECALL OF PROSE

BY MICHAEL W. EYSENCK
Department of Psychology, Birkbeck College, University of London

Subjects were divided into four groups on the basis of their scores on the extraversion scale of the Eysenck Personality Inventory and the general activation scale of Thayer's Activation–Deactivation Adjective Check List. A prose passage was presented to subjects who then attempted to recall it under strict instructions to avoid errors. There were significant interactions between activation and extraversion for the error data and the phrase-correct data, with moderate levels of arousal (high activation extraverts and low activation introverts) being associated with the fewest errors and the most phrases correctly recalled. It was concluded that the study showed the importance of arousal to recall performance, possibly due to the effects of arousal on retrieval processes.

Several studies have looked at the relationship between extraversion and immediate recall. A frequent finding has been that extraverts recall significantly more than introverts, particularly under conditions involving some distraction or response interference (Allsopp & H. J. Eysenck, 1975; Bone, 1971; Howarth, 1969*a*, *b*; Jensen, 1964; Shanmugan & Santhanam, 1964; Siegman, 1957). Theoretically, a number of these investigators assumed that the inferior recall of introverts was due to their high level of arousal. Experimental evidence that introverts are more highly aroused than extraverts was originally derived from studies comparing introverts and extraverts on the EEG (Gale *et al.*, 1969; Savage, 1964). Subsequent work has been reviewed by Gale (1973), who cites seven experiments obtaining evidence that introverts are, indeed, more highly aroused on EEG indices than are extraverts. Furthermore, a considerable body of behavioural evidence in the fields of motivation, conditioning, perception, psychopharmacology, vigilance and learning is consistent with the assumption that introverts are characterized by a high level of arousal (H. J. Eysenck, 1967). Evidence that both low and high levels of subject arousal are detrimental to immediate recall comes in studies by Allsopp & H. J. Eysenck (1974), M. W. Eysenck (1975*b*) and McLaughlin & H. J. Eysenck (1967).

A distinction has sometimes (e.g. H. J. Eysenck, 1967) been drawn between arousal, seen as arising from activity of the reticular formation, and activation, defined as emotional arousal, and considered to be a product of the activity of the hypothalamic limbic system. For present purposes, it will be assumed that 'arousal' refers to some elevated state of bodily function, that it represents a non-specific increment in physiological activity, and that it is non-informational. In common with the research literature on arousal and learning (see Levonian, 1972; Uehling, 1972, for reviews), the simplifying assumption will be made that arousal is a unitary concept. Ultimately, however, it is possible that a separation of arousal into electrocortical arousal, autonomic arousal and behavioural arousal, as recommended by Lacey (1967), will be necessary.

An interpretative difficulty with the above studies is that they involved the learning of lists of individual words or paired associates. In such studies, the incidence of errors is low, possibly due to the operation of an efficient editing process

(Peterson, 1967), and statistical analyses of the error data were not carried out. Thus it is difficult to decide whether the inferior recall of high-arousal subjects might merely reflect a greater cautiousness in responding. One of the aims of this study was to investigate the effects of subject arousal on recall in a task generating numerous errors. Tales taken from the North American Indian culture appeared suitable in this context (Bartlett, 1932). Poor recall by high-arousal subjects can be attributed, at least in part, to cautious responding if high-arousal subjects make relatively few errors, but is more plausibly attributed to inadequate storage or retrieval if high arousal is associated with a relatively high error rate. A second aim of this study was to investigate the extent to which the performance differences found in verbal learning as a function of subject arousal are obtainable in the context of a study on the recall of connected discourse.

A consideration of individual differences in correct recall and in incidence of errors is relevant to the theoretical positions of Bartlett (1932) and of Zangwill (1972). Bartlett's (1932) constructive theory emphasized the importance of the errors and distortions which subjects introduce in their attempted recall of connected discourse, whereas Gomulicki (1956*a*) and Zangwill (1972) stressed the abstractive nature of recall. They suggested that recall consists of an abstract of the original passage, similar in many ways to an intentional abstract or précis. The possibility that individual differences may play a part in determining whether recall is constructive or abstractive was indicated in a study by Gomulicki (1956*b*). He read short passages to his subjects, and asked them to repeat orally as much as they could recall immediately after listening to each passage. His major finding was that subjects' scores were bimodally distributed with respect to a relative preference for changes or for omissions.

Previous studies investigating the effects of level of subject arousal on learning (e.g. Allsopp & H. J. Eysenck, 1974, 1975) have considered arousal only in the sense of a characteristic, semi-permanent, level of arousal. As the physiological evidence makes clear, 'arousal' can also refer to the subject's transient level of arousal within the experimental situation (cf. Uehling, 1972). In order to obtain as complete an assessment of the relationship between subject arousal and learning as possible, characteristic levels of arousal were measured by the Eysenck Personality Inventory (EPI), and transient levels of arousal were indexed by the General Activation scale of Thayer's Activation–Deactivation Adjective Check List. This latter test, which is an objective self-report measure of transient levels of arousal, was investigated in several experiments by Thayer (1967, 1970). His main finding was that general activation showed substantial correlations with physiological measures of arousal.

A methodological consideration of some consequence was identified by Gauld & Stephenson (1967). They found that many of the errors and distortions in recall obtained by Bartlett's (1932) subjects could be attributed to deliberate inventions of material to fill in gaps in their memories rather than to genuine memorial distortion. An attempt was made to see whether the various groups of subjects differed in the extent to which they indulged in confabulation by using two separate recall indices.

METHOD

Subjects

The subjects were 64 college students. They were divided into four equal-sized groups on the basis of their scores on the extraversion scale of the EPI (Form 1), and the general activation scale of the short-form version of Thayer's (1967) Activation–Deactivation Adjective Check List. All those scoring 11 on the extraversion scale, 4 or more on the lie scale, or 12 on the general activation scale, were excluded. The four groups represented all the possible combinations of high- and low-extraversion and high- and low-activation. The groups did not differ in their mean neuroticism scores on the EPI. The means and standard deviations of the various scales were as follows: EPI–extraversion scale: 10·92 (4·04); EPI–neuroticism scale: 9·41 (5·17); EPI–lie scale: 1·41 (1·12); and general activation scale: 11·31 (3·46).

Materials and procedure

Subjects initially completed the Activation–Deactivation Adjective Check List and the EPI. They were then given the following instructions: 'Listen carefully to the following story. You will be asked later on to reproduce it. Please have paper and pen ready.'

Subjects were provided with a typewritten copy of 'The War of the Ghosts' (Bartlett, 1932). The male experimenter read the story aloud, with the subjects following the story in their texts. The typewritten copies were turned over and collected at the conclusion of the story, followed by the experimenter giving the instructions: 'I want you to set down on paper as much as you can remember about the story I have just read you. It is particularly important that you don't set down facts and incidents which were not in the original. If there are parts of the original which you just do not remember, do not invent material to fill the gaps for the sake of being able to tell a complete story: leave blanks. Please start now.'

Fifteen minutes was allowed for written recall. At the end of the time period, the experimenter gave the subjects further instructions: 'Please look through your version of the story incident by incident and fact by fact, and ask yourself whether or not you're quite certain that all the facts and incidents you set down were really in the original story. Do not indicate that you are certain if there is any doubt. Please enclose any phrases of which you are not certain in square brackets.' Unlimited time was allowed for this second recall test. In sum, the basic design was a 2 (activation) × 2 (extraversion) × 2 (recall test) analysis of variance design, with the first two factors between subjects and the last factor within subjects.

Measurement of performance

In order to have a measure of error tendency that was unconfounded by the number of words recalled, the error-rate index of Gauld & Stephenson (1967) was employed. This index was defined as the mean number of errors per 100 words of recall. An error was considered to be any deviation whatsoever from the original passage except for synonyms or synonymous phrases; omissions; time order errors; place name errors. Furthermore, if an error was repeated, it was scored as an error only once.

The number of correct phrases recalled was based upon a division of the original passage into 60 phrases. Any phrase correctly recalled or any phrase synonymous with a phrase from the original passage scored one; any phrase partially recalled but omitting an important aspect of the phrase from the original passage scored one-half. Both this measure and the error-rate index were calculated by three judges, and the mean taken.

RESULTS

An initial analysis of the data was concerned with the number of words recalled on each of the two recall tests. A 2 (activation) × 2 (extraversion) × 2 (recall test) analysis of variance, with the first two factors between subjects, and the last factor within subjects, was performed on the data of Table 1. As anticipated, significantly more words were recalled on the first test than on the second. The main effect of

Table 1. *Total words recalled as a function of extraversion and activation: data and analysis*

Extraversion ...	Low	Low	High	High
Activation ...	Low	High	Low	High
Recall test 1				
Mean	190·06	167·88	179·75	208·12
S.D.	23·68	48·08	23·70	36·05
Recall test 2				
Mean	184·00	155·75	163·63	193·81
S.D.	25·10	48·60	27·33	35·44

Source	D.F.	M.S.	*F*
Between subjects	63		
Activation (A)	1	132·03	< 1
Extraversion (C)	1	4536·28	1·93
AC	1	23762·00	10·13**
Subjects within groups	60	2345·96	
Within subjects	64		
Recall test (B)	1	4728·78	55·04***
AB	1	36·13	< 1
BC	1	300·13	3·49
ABC	1	124·03	1·44
B × subjects within groups	60	85·92	
Total	127		

* $P = 0·05$; ** $P = 0·01$; *** $P = 0·001$.

activation was not significant, and neither was the main effect of extraversion. However, there was a significant interaction between activation and extraversion, as indicated in Table 1. In this interaction, high-activation extraverts recalled more words than low-activation extraverts, but high-activation introverts recalled fewer words than low-activation introverts. None of the remaining interactions was significant.

A second three-way analysis of variance was conducted on the correct-phrase data presented in Table 2. As can be seen from reference to Table 2, the only significant main effect was that of recall test, with significantly more phrases being correctly recalled on the first recall test than on the second. The only significant interaction was between extraversion and activation. In this interaction, shown in Fig. 1, the group with the lowest level of arousal (low activation extraverts) and the group with the highest level of arousal (high activation introverts) recalled the fewest phrases, and the groups of intermediate arousal (high activation extraverts and low activation introverts) recalled the most phrases.

A third three-way analysis of variance was performed on the error-rate data presented in Table 3. Gauld & Stephenson (1967) found that the error rate was significantly lower on the corrected version of the written reproduction than on the initial recall, and this finding was replicated in the current study. There was a highly significant difference in error rate between the two recalls, as indicated in Table 3. In addition, the main effect of activation was significant, with high activation being associated with a high error rate. Introverts tended to have a higher error rate than extraverts, but the difference did not attain the conventional level of statistical significance. The interaction between activation and extraversion, which was significant in the previous two analyses, was marginally significant. As can be seen in

Table 2. *Number of phrases correctly recalled as a function of extraversion and activation: data and analysis*

Extraversion ...	Low	Low	High	High
Activation ...	Low	High	Low	High
Recall test 1				
Mean	24·75	20·94	22·13	26·50
S.D.	5·01	6·66	2·50	6·13
Recall test 2				
Mean	24·19	20·38	20·94	25·38
S.D.	4 90	6·73	2·64	6·38

Source	D.F.	M.S.	F
Between subjects	63		
Activation (A)	1	2·82	< 1
Extraversion (C)	1	43·95	< 1
AC	1	540·38	9·44**
Subjects within groups	60	57·24	
Within subjects	64		
Recall test	1	23·63	71·61***
AB	1	0·01	< 1
BC	1	2·82	< 1
ABC	1	0·01	< 1
B × subjects within groups	60	0·33	
Total	127		

* $P = 0·05$; ** $P = 0·01$; *** $P = 0·001$.

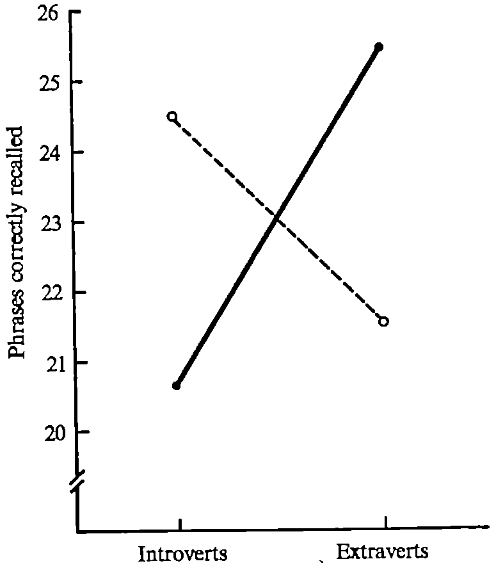


Fig. 1. Phrases correctly recalled as a function of activation and of introversion–extraversion. ●, High activation; ○, low activation.

Table 3. *Error rate as a function of extraversion and activation: data and analysis*

Extraversion ...	Low	Low	High	High
Activation ...	Low	High	Low	High
Recall test 1				
Mean	2·46	3·90	2·57	2·92
S.D.	1·18	1·42	0·69	1·00
Recall test 2				
Mean	2·30	3·60	2·23	2·65
S.D.	1·16	1·55	0·78	0·71

Source	D.F.	M.S.	*F*
Between subjects	63		
Activation (A)	1	247016·63	10·62**
Extraversion (C)	1	71489·26	3·07
AC	1	78061·88	3·36
Subjects within groups	60	23260·16	
Within subjects	64		
Recall test (B)	1	22128·82	21·75**
AB	1	89·45	< 1
BC	1	476·63	< 1
ABC	1	929·88	< 1
B × subjects within groups	60	1017·28	
Total	127		

* $P = 0·05$; ** $P = 0·01$; *** $P = 0·001$.

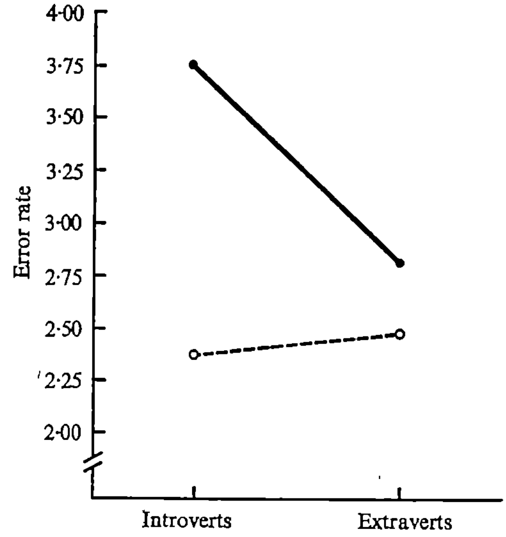


Fig. 2. Error rate as a function of activation and of introversion–extraversion. ●, High activation; ○, low activation.

Fig. 2, high activation again had a more detrimental effect on introverts than on extraverts.

A final analysis correlated the correct-phrase data and the error-rate data for each of the recall tests separately. The Spearman rank-order correlation coefficient was $-0·39$ ($P < 0·01$) for the data of the first recall, and $-0·37$ ($P < 0·01$) for the data of the corrected recall. In other words, those subjects with high levels of recall tended to have a lower error rate than those who recalled less.

DISCUSSION

The finding that there was a curvilinear relationship between subject arousal and the number of phrases recalled corresponds to results obtained in studies of paired-associate learning (Allsopp & H. J. Eysenck, 1974; M. W. Eysenck, 1975*b*; McLaughlin & H. J. Eysenck, 1967). The relationship thus has a certain generality and is not merely paradigm-specific. It is worth emphasizing this generality, given the prevalent assumption that findings of general applicability are unlikely to emerge from the study of individual differences. For example, Cofer (1967) has stated: 'I remain to be convinced...that there are but a few basic processes underlying individual differences in learning, and that the interactions of specific task characteristics with highly specific individual propensities are not the basis of the enormous individual differences which everyone finds in most learning tasks' (p. 139).

Any attempt to explain the relatively poor recall performance of subjects low and high in arousal must consider at least three major alternative hypotheses. The first hypothesis is simply that the initial registration and storage of the passage has been inadequate. However, several studies have shown that, while high-arousal subjects have inferior recall to low-arousal subjects on an immediate recall test, the reverse is the case at longer retention intervals (e.g. Howarth & H. J. Eysenck, 1968; McLean, 1968; Osborne, 1972; Skanthakumari, 1965). The high levels of recall attained by high-arousal on delayed retention tests argues against this hypothesis.

The second hypothesis is that subjects of low and high arousal have difficulty in retrieving the information to be remembered. A recent series of studies has investigated this hypothesis (M. W. Eysenck, 1974*a*, *b*, 1975*a*, *b*). In all of these studies, attempts were made to equate the various groups in terms of storage. Two crucial findings were obtained: (1) subjects of low and high arousal recalled less than those of intermediate arousal, even when lengthy retrieval periods were available (M. W. Eysenck, 1974*a*); (2) low- and high-arousal subjects took longer than those of intermediate arousal to retrieve information (M. W. Eysenck, 1974*b*, 1975*a*, *b*). The evidence thus suggests that at least part of the reason for the poor recall performance of low- and high-arousal subjects derives from difficulties of search and retrieval.

The third hypothesis assumes that subjects low and high in arousal retrieve information as well as those of intermediate arousal, but that they are more cautious in their responding. While M. W. Eysenck (1974*b*) obtained some evidence against this caution hypothesis, it is the case that the long recall latencies of low and high arousal subjects could be due either to inefficient search processes or to a higher threshold for response.

The results of this study provide relevant information for investigating this third hypothesis. At least two testable predictions can be made from this hypothesis. The first is that subjects low and high in arousal should have lower error scores than subjects intermediate in arousal. The second is that, for the error-rate data, there should be a three-way interaction among activation, extraversion, and recall test. On the caution hypothesis, the more stringent second recall test should have led to a greater reduction in error rate for subjects of intermediate arousal than for those low or high in arousal. The reason for this is that the low-activation extraverts and the high-activation introverts are already adopting cautious criteria on the first recall test. Neither of these predictions was supported. High-activation introverts had the highest error rate, whereas they were predicted to have a low error rate. Although there was a highly significant reduction in the error rate on the second recall test, the recall test factor did not interact significantly with activation or extraversion either individually or jointly (all F's < 1).

The desirability of considering arousal both as a semi-permanent aspect of individual differences and as a transient state was demonstrated. In none of the three major analyses was there a significant main effect of extraversion, whereas the interaction between extraversion and activation was highly significant in two of the analyses, and of marginal significance in the third. It is probable that factor theories of personality over-emphasize the stability and durability of personality characteristics.

Clear support was obtained for Gauld & Stephenson's (1967) contention that the errors in reproduction found by Bartlett (1932) included much deliberate confabulation. The instructions given prior to the first recall test required the subjects to adopt a considerably more stringent criterion than the instructions used by Bartlett (1932), and still the subjects showed some ability to distinguish between those parts of their recall which corresponded to the original passage and those which did not.

Gomulicki (1956*a*) found a tendency for subjects either to make several errors in their recall or to omit a substantial proportion of the passage. He did not investigate the personality or other differences between these two types of subjects. In this study, in contrast, there was a significant tendency for those subjects who made several errors to be the subjects who also omitted most from the passage. More specifically, the high-activation introverts had a disproportionate number of both commission and omission errors.

REFERENCES

ALLSOPP, J. F. & EYSENCK, H. J. (1974). Personality as a determinant of paired-associates learning. *Percept. mot. Skills* **39**, 315–324.

ALLSOPP, J. F. & EYSENCK, H. J. (1975). Extraversion, neuroticism and verbal reasoning ability as determinants of paired-associates learning. *Br. J. Psychol.* **66**, 15–24.

BARTLETT, F. C. (1932). *Remembering.* Cambridge University Press.

BONE, R. N. (1971). Interference, extraversion and paired-associate learning. *Br. J. soc. clin. Psychol.* **10**, 284–285.

COFER, C. N. (1967). The specificity of individual differences in learning. In R. M. Gagné (ed.). *Learning and Individual Differences.* Columbus, Ohio: Merrill.

EYSENCK, H. J. (1967). *The Biological Basis of Personality.* Springfield, Ill.: Thomas.

EYSENCK, M. W. (1974*a*). Extraversion, arousal and retrieval from semantic memory. *J. Personality* **42**, 319–331.

EYSENCK, M. W. (1974*b*). Individual differences in speed of retrieval from semantic memory. *J. exp. Res. Person.* **8**, 307–323.

EYSENCK, M. W. (1975*a*). Effects of noise, activation level, and response dominance on retrieval from semantic memory. *J. exp. Psychol.* **104**, 143–148.

EYSENCK, M. W. (1975*b*). Arousal and speed of recall. *Br. J. soc. clin. Psychol.* **14**, 269–277.

GALE, A. (1973). The psychophysiology of individual differences: studies of extraversion and the EEG. In P. Kline (ed.), *New Approaches to Psychological Measurement.* London: Wiley.

GALE, A., COLES, M. & BLAYDON, J. (1969). Extraversion–introversion and the EEG. *Br. J. Psychol.* **60**, 209–223.

GAULD, A. & STEPHENSON, G. M. (1967). Some experiments relating to Bartlett's theory of remembering. *Br. J. Psychol.* **58**, 39–49.

GOMULICKI, B. R. (1956*a*). Individual differences in recall. *J. Personality* **24**, 387–400.

GOMULICKI, B. R. (1956*b*). Recall as an abstractive process. *Acta psychol.* **12**, 77–94.

HOWARTH, E. (1969*a*). Personality differences in serial learning under distraction. *Percept. mot. Skills* **28**, 379–382.

HOWARTH, E. (1969*b*). Extraversion and increased interference in paired-associate learning. *Percept. mot. Skills* **29**, 403–406.

HOWARTH, E. & EYSENCK, H. J. (1968). Extraversion, arousal, and paired-associate recall. *J. exp. Res. Person.* **3**, 114–116.

JENSEN, A. (1964). Individual differences in learning: interference factor. (Cooperative Research Project no. 1867, Office of Education, U.S. Department of Health, Education and Welfare.)

LACEY, J. I. (1967). Somatic response patterning and stress: some revisions of activation theory. In M. H. Appley & R. Trumbull (eds.), *Psychological Stress*. New York: Appleton-Century-Crofts.

LEVONIAN, E. (1972). Retention over time in relation to arousal during learning: an explanation of discrepant results. *Acta psychol.* **36**, 290–321.

MCLAUGHLIN, R. J. & EYSENCK, H. J. (1967). Extraversion, neuroticism, and paired-associates learning. *J. exp. Res. Person.* **2**, 128–132.

OSBORNE, J. W. (1972). Short- and long-term memory as a function of individual differences in arousal. *Percept. mot. Skills* **34**, 587–593,

PETERSON, L. R. (1967). Search and judgment in memory. In B. Kleinmuntz (ed.), *Concepts and the Structure of Memory*. London: Wiley.

SAVAGE, R. D. (1964). Electrocerebral activity, extraversion and neuroticism. *Br. J. Psychiat.* **110**, 98–100.

SHANMUGAN, T. E. & SANTHANAM, M. C. (1964). Personality differences in serial learning when interference is presented at the marginal visual level. *J. Ind. Acad. Appl. Psychol.* **1**, 25–28.

SIEGMAN, A. W. (1957). Some relationships of anxiety and introversion–extraversion to serial learning. (Unpublished Ph.D. thesis, Ann Arbor, Michigan.)

SKANTHAKUMARI, S. R. (1965). Personality differences in the rate of forgetting. *J. Ind. Acad. Appl. Psychol.* **2**, 39–47.

THAYER, R. E. (1967). Measurement of activation through self-report. *Psychol. Rep.* **20**, 663–678.

THAYER, R. E. (1970). Activation states as assessed by verbal report and four psychophysiological variables. *Psychophysiol.* **7**, 86–94.

UEHLING, B. S. (1972). Arousal in verbal learning. In *Human Memory: Festschrift in Honour of Benton J. Underwood*. New York: Appleton-Century-Crofts.

ZANGWILL, O. L. (1972). 'Remembering' revisited. *Q. Jl exp. Psychol.* **24**, 123–138.

(*Manuscript received* 4 *February* 1975; *revised manuscript received* 1 *April* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 63–72
Printed in Great Britain

THE RECALL OF CONCRETE AND ABSTRACT SENTENCES AS A FUNCTION OF INTERPOLATED TASK

BY GRAHAM DAVIES AND JOHN PROCTOR
Department of Psychology, University of Aberdeen

Begg & Paivio (1969) have proposed that while sentences dealing with abstract matters are normally coded in verbal terms, those concerned with concrete events are internalized as imagery. In two experiments, subjects observed both types of sentence prior to carrying out a distracter task which involved either verbal or perceptual coding. In agreement with the hypothesis, sentence type interacted significantly with interpolated task in both studies. Abstract sentence recall was significantly more impaired following the verbal task than the perceptual, while this effect was reversed for concrete sentences. In the second experiment, this latter difference also reached significance and, further, error analysis confirmed that the perceptual task influenced the semantic content of the sentences recalled. The results are discussed in terms of Begg & Paivio's hypothesis. Certain extensions and limitations of the theory are then considered.

Until recently, theories of how sentences are memorized have emphasized the purely linguistic aspects of the coding process (e.g. Miller, 1962; Mehler, 1963). However, Begg & Paivio (1969) drew attention to the possibility that both verbal and non-verbal processes might mediate recall. Sentences concerned with concrete events ('The fat girl hit the boy') might normally be encoded in terms of a unitary mental image. Recall would then involve retrieval of the image and reconstruction of the original sentence on the basis of the contents. Abstract sentences on the other hand ('A good religion teaches charity') would normally be remembered as a sequentially organized string of words.

Probably the best evidence in support of this proposition has come from a study by Klee & Eysenck (1973) on sentence comprehension. They asked subjects to read a series of concrete and abstract sentences and then to verify whether or not they made sense. Interpolated between words in each sentence was a task designed to interfere selectively with either perceptual or verbal coding. Perceptual interference was provided by asking the subjects to observe a series of matrices which they later had to reproduce. Verbal interference was occasioned by requiring subjects to read a series of digits which had to be remembered afterwards. For concrete sentences, the perceptual task led to longer comprehension latencies than did the verbal. Consistent with the Begg & Paivio hypothesis, when abstract sentences were employed the relative interfering effects of the two tasks was reversed.

As yet, however, attempts to provide direct evidence for the hypothesis from studies of sentence retention have met only with mixed success. The results of Begg & Paivio's own study on recognition memory have been effectively criticized by Johnson *et al.* (1972) as being largely attributable to comprehension factors rather than memory as such. Sasson (1971) and Sasson & Fraisse (1972) followed exposure to concrete and abstract sentences with the rapid presentation of either further sentences or pictures designed to produce selective interference effects. Although the results were in the expected direction, pictures did not produce significantly greater loss of recall for the concrete sentences than did the verbal material, nor was the effect significantly reversed in the abstract condition.

Atwood (1971), on the other hand, did report the expected interaction between sentence type and the form of interpolated task. However, his subjects were specifically requested to form an image for the concrete sentences and 'to contemplate the meaning' of the abstract. Thus the results may be of limited value as a guide as to how subjects normally encode such material. Further, repeated attempts by Brooks and others (see Paivio, 1971, p. 374) to replicate the original findings have not been successful.

A possible explanation for the anomalous findings of these memory studies may lie in the nature of the interpolated tasks employed. The rate of presentation employed by Sasson for the picture material in his perceptual interference condition was quite sufficient for the subject to verbally label isolated features (Paivio & Csapo, 1969). Equally, the time allowed to view the interfering sentences in his verbal condition has been shown to be long enough to enable many subjects to generate at least some visual imagery (Paivio & Begg, 1971). In Atwood's study, both interference conditions involved the presentation of either the number '1' or '2' to which the subject responded by saying the opposite. In the perceptual interference condition the first number was presented visually, whereas in the verbal condition, auditory presentation was employed. Once again there appears to be the possibility of contamination of the perceptual interference task by verbal elements. Not only is verbal coding involved in the processing of the initial stimulus but also in the subject's reponse. As Waugh & Norman (1965) demonstrated, verbal output produces as much loss of information from short-term memory as verbal input.

The present study seeks to overcome the difficulties associated with these earlier recall experiments through the use of interpolated tasks designed to produce purer measures of the two forms of interference. Verbal interference was produced by asking the subject to count backwards in single numbers. This was judged to be sufficiently complex as to prevent rehearsal of the target sentences yet not so taxing as to produce the elaborate mental calculating noted by Meudell (1972) with more complex subtraction tasks. The perceptual task was an adaptation of the tracking procedure first employed as an interference technique by Bower (1972). The subject was required to maintain a stylus within the boundaries delineated by two parallel lines drawn on a paper roll which unwound in front of him.

EXPERIMENT I

In this first study, subjects' ability to retain concrete abstract sentences was compared under conditions designed to produce interference to either verbally or perceptually stored information. It was predicted that sentence type would interact with interpolated task such that recall of abstract sentences would be more impaired by verbal interference while concrete sentences would suffer more from perceptual interference.

Method

Subjects

The subjects were 44 female and 16 male members of the Departmental Subject Panel. Their ages ranged from 16 to 45 years and they were drawn from a wide range of social and educational backgrounds. They were allocated to one of four groups on the basis of age and sex.

Design

Two groups of subjects observed three concrete sentences while the other two looked at a corresponding set of abstract sentences. Following this, the groups performed either on a verbal or perceptual interfering task before recalling the sentences. Crossing sentence type with form of interpolated task produced four conditions corresponding to the four groups.

Selection of sentences

In order to ensure strict comparability between concrete and abstract sentences, matched pairs of sentences were generated where each member of the pair had an identical structure. However, in one version all the nouns were concrete, while in the other they were abstract. The choice of nouns was guided by the concreteness norms of Paivio *et al.* (1968), supplemented by unpublished normative data collected by the first author. Frequency and meaningfulness (*m*) was equated across members of a given sentence pair. Each sentence contained three nouns connected by a single verb. The 15 most plausible and sensible sets of sentences so generated were typed in a random order onto a sheet of paper and distributed to a separate group of 20 comparable members of the subject panel. They were requested, following a procedure adopted from Sutcliffe (1969), to rate each sentence on its image-evoking capacity. A seven-point scale was used ranging from '1' ('perfectly clear and as vivid as the actual experience') through to '7' ('no image present at all, you only knowing you are thinking of the concept'). This two stage selection process was employed in view of the possible ameliorating effects of context on concreteness ratings (Morris & Reid, 1972). The three pairs of sentences which produced the largest differences in mean ratings were selected as stimulus materials. The pairs of sentences employed are printed below. The mean imagery ratings of the concrete and abstract versions are shown in parentheses:

Wolves follow *cattle* and *sheep* (1·50)
Memories follow *glory* and *honour* (4·30)

The *flask* of *lemonade* is in fact a *bomb* (2·35)
The *concept* of *greed* is in fact a *fallacy* (5·45)

The *soldier* and the *car* meet up in the *mountains* (2·20)
The *soul* and the *spirit* meet up in the *heavens* (5·10)

Procedure

All subjects were tested individually and in a random order with regard to condition. The procedure to be followed was outlined to the subjects. They were told they were to be shown three sentences which they were to read aloud and 'attempt to understand'. They were also warned that they would have to provide a verbatim report of the sentences immediately following a short interpolated task, the nature of which was explained to them. Apart from this, no special instructions were given as to the mental set to be adopted in learning the material. Once the experimenter was satisfied that the instructions were understood, he gave the subject a pack of three file cards with a different sentence typed on each. The subject took each card in turn and read aloud the sentence printed on it. The order in which the sentences were read was varied systematically within conditions according to a latin square design. Immediately following observation of the final sentence, the subject took part in the appropriate interference task for 20 sec.

In the *verbal* task, the experimenter announced a three figure number at random and the subject was required to count backward from it until told to stop.

In the *perceptual* condition, the subject was handed a stylus attached to a tracking machine by an electrical lead. The track consisted of two parallel lines approximately 1½ cm apart drawn on a standard paper tape 8½ cm wide. In order to make the task difficult and challenging, the

Table 1. *Mean number of nouns recalled as a function of type of sentence and form of interpolated task*

(Mean complete sentences recalled in brackets.)

	Type of interpolated task	
Sentences	Verbal	Perceptual
(*a*) Experiment I		
Abstract		
Mean	2·80	5·27
S.D.	1·75	1·83
	(0·40)	(0·93)
Concrete		
Mean	7·13	6·53
S.D.	1·51	1·75
	(1·60)	(1·13)
(*b*) Experiment II		
Abstract		
Mean	3·17	4·28
S.D.	1·41	1·93
	(0·17)	(0·66)
Concrete		
Mean	8·22	7·05
S.D.	0·73	1·20
	(2·22)	(1·33)

track veered erratically across the tape. The tape unwound slowly behind the screen in which a small slit had been cut. This enabled the subject to view a 3 cm section of track at any one time. The subjects were instructed to attempt to keep the stylus head in contact with the paper and within the boundaries set by the lines. In order to ensure maximum attention to the interpolated task and the minimum of verbal rehearsal, subjects were informed that their tracking performance was being monitored by an impressive but totally non-functional piece of apparatus placed at the experimenter's elbow. Subjects in the perceptual conditions were allowed a short familiarizing period on the tracking apparatus prior to the experimental trial.

At the end of 20 sec. all subjects were stopped and asked to perform a cued recall test for the words in the sentences they had observed previously. The cue cards, which were the same for both concrete and abstract conditions, showed the three original sentences minus their main nouns. Subjects were requested to say aloud as much of the target sentences as they could remember. The recall trial was unpaced and continued until the subject had satisfied himself that he could remember no more. In no instance did this take longer than two minutes. All responses made by the subject were recorded on tape for later analysis.

Scoring

Gross performance was scored by awarding subjects one point for each noun correctly recalled, irrespective of position. As there were three nouns in each sentence, every subject was allotted a score out of 9. A supplementary analysis in terms of total number of sentences recalled entirely correctly was also performed to ensure the reliability of the first procedure.

Results and discussion

The average number of nouns recalled correctly as a function of condition of presentation and form of interpolated task is shown in Table 1 (*a*). The recall scores were analysed by means of a 2 × 2 analysis of variance with sentence type and interpolated task as main effects. This indicated while concrete sentences were better recalled than abstract ($F = 25{\cdot}11$; d.f. $= 1,56$; $P < 0{\cdot}001$), this effect was qualified by the predicted interaction between sentence type and form of interpolated task ($F = 5{\cdot}07$; d.f. $= 1,56$; $P < 0{\cdot}05$). Examination of the simple effects underlying this interaction (Kirk, 1968) indicated that when abstract sentences were presented, the perceptual task interfered less with recall than the verbal task ($P < 0{\cdot}05$). When, on the other hand, concrete sentences were employed, the perceptual task produced an insignificantly greater number of errors than the verbal. The main effects for the form of the interference task were negligible ($F < 1$), indicating that the tracking and counting tasks were of approximately equal levels of overall difficulty. Examination of the mean number of sentences recalled by each subject as a function of condition (see Table 1*a*) confirmed the findings of the main analysis. Thus the main results of the present study offer qualified support for the Begg & Paivio hypothesis. As predicted perceptual tracking does produce less interference in the recall of abstract sentences than the verbal task, but the predicted reversal of the effect with concrete sentences, though in the expected direction, does not reach significance.

An examination was also conducted into the types of error produced by subjects in the concrete and abstract conditions. The bulk of errors were simple omissions rather than intrusions from outside the list. Omissions were more common in the abstract conditions (93 per cent of all errors) than in the concrete (62 per cent). Transposition errors, transposing nouns from one sentence to another, only occurred for abstract sentences. In all, seven of the 30 subjects in the abstract conditions showed such errors. The form of the interpolated task appeared to have no marked effect on the relative frequency of error types.

Additional analyses were also performed to examine any effects due to order of presentation and the individual sentences employed. No significant effects for sentence order were found nor was this variable influenced by either the concreteness of the sentences concerned or the nature of the interpolated task (all F's < 1). The sentences themselves, however, differ in their individual memorability both between sentence frames ($F = 4{\cdot}88$; d.f. $= 2, 112$; $P < 0{\cdot}025$) and within the concrete and abstract sets ($F = 6{\cdot}65$; d.f. $= 2,112$; $P < 0{\cdot}01$). However, the effects of the interference tasks were uniform across sentences.

Clearly, while these initial results appear promising, they are insufficient in themselves to provide convincing support for the Paivio & Begg hypothesis. In particular, the failure of the two interference tasks to have differential effects for concrete sentences was disappointing as was the large measure of variability shown both across sentences and between individuals. Accordingly a second experiment was undertaken which incorporated a number of modifications designed to amplify the results of the initial study.

Experiment II

In this second study, inter-subject variability was more closely controlled by the use of a more homogeneous subject population and by incorporating the effects of sex in the main analysis. A number of minor changes in procedure were also undertaken in an attempt to clarify the effects in the concrete sentence conditions. The most important of these was the decision to extend the length of time spent on the interpolated task in the hope that the resulting increased degradation in the trace would accentuate any differences in performance due to the interference procedures.

Method

Subjects

The subjects for this second study were 36 male and 36 female volunteers, aged 18–26 years, drawn from the undergraduate population and clerical staff of the University of Aberdeen. They were randomly assigned to one of four groups with the constraint that equal numbers of each sex should be represented in all groups.

Selection of sentences

The same procedure for the selection of stimulus material was employed as in Expt. I. Pairs of concrete and abstract sentences, shown in a random order, were submitted to an independent panel of 40 student judges who rated them for imagery content. The three pairs of sentences showing the largest disparity in rating were selected for use and are shown below, together with their mean ratings on the Sutcliffe scale:

The *doorman* of the *emporium* was a *wrestler* (2·84)
The *disclosure* of the *folly* was an *indiscretion* (5·18)

The *shoes* are a product of the *skin* of the *cow* (2·84)
The *anxieties* are a product of the *fear* of the *unknown* (5·56)

The *attic* contains great *trunks* and also some *paintings* (1·97)
The *belief* contains great *wisdom* and also some *illusions* (5·57)

Design and procedure

The design of the present study was identical save for the inclusion of sex of subject as a formal variable. The procedure, too, followed the same pattern except in two respects. First, the time spent on the interpolated tasks was extended from 20 to 30 sec. Second, the individual sentences were no longer submitted to the subject on separate cards but appeared together on one card. This latter change was an attempt, by eliminating movement time between sentences, to reduce still further the possibility of rehearsal.

Results and discussion

As Table 1(*b*) makes clear, the minor changes in experimental format had the desired effect of clarifying the pattern of the original results. A $2 \times 2 \times 2$ analysis of variance was performed on the number of nouns recalled by subjects as a function of sex, type of sentence and form of interpolated task. Concrete sentences were once again better recalled than abstract ($F = 129{\cdot}03$; d.f. $= 1, 64$; $P < 0{\cdot}001$), while the interference procedures had much the same overall effect on performance ($F < 1$). The interaction between sentence type and interpolated task was again replicated, but at a higher level of significance ($F = 10{\cdot}91$; d.f. $= 1, 64$; $P < 0{\cdot}01$). A simple effects analysis confirmed that when abstract sentences were observed, counting led to poorer recall than tracking ($P < 0{\cdot}001$). More importantly, however, the effect was significantly reversed for concrete sentences ($P < 0{\cdot}001$). None of these effects was in any way qualified by the sex of the subject nor did the main effects of this variable approach significance ($F < 1$). Further, all these findings were confirmed by supplementary analysis in terms of total complete sentences recalled.

Error analysis again showed differences between the concrete and abstract conditions. Once more, transposition errors were confined to abstract sentences with 10 out of 30 subjects showing such errors. However, the relative frequency of intrusions and omissions with the two types of sentence differed from that in Expt. I. Consistent with the previous results, omissions formed the vast majority of all errors on the abstract conditions (84 per cent), but the pattern was completely reversed in the concrete conditions. On this occasion, subjects never failed to provide a response when called upon to fill a blank in a concrete sentence. In signal detection theory terms (see Banks, 1970), the students in Expt. II operated on a laxer criterion than the subject panel members employed previously. Consistent with the hypothesis, the intrusions produced after the tracking task betrayed more serious semantic distortions than those induced by backward counting. Two independent judges divided the concrete sentence intrusions into those reflecting a change of meaning from the original and those merely showing a change of wording. Their analysis showed a significantly higher rate of meaning changes following the perceptual task (77 per cent of errors) than after the verbal (50 per cent).

The additional analyses of the effects of presentation order and sentences employed followed the same pattern as the previous experiment. There were no significant effects due to presentation order but sentences again differed in their ease of memorability. On this occasion the latter effect was confined to differences between the sentence frames ($F = 14{\cdot}57$; d.f. $= 2, 136$; $P < 0{\cdot}001$), the relative memorability of the concrete and abstract versions of a given sentence frame being much the same. As before, neither order of presentation nor sentence frame had any influence on the effects of the interpolated task.

Thus the present results provide additional support for the main hypothesis. The minor changes in procedure and the use of a more selected sample of subjects appear to have had the desired effect of clarifying and increasing the selective interference effects with sentences of differing imagery value, noted in the earlier study.

GENERAL DISCUSSION

An interaction between sentence type and form of interference task has now been obtained in two studies, using different samples of subjects and ranges of experimental material. In both instances the difference in performance on abstract sentences following tracking and counting has been significant. The reversal of effect with concrete sentences reached significance when greater time was expended on the interpolated tasks. Taken together, the results of these studies appear to provide strong support for the view that the mediating processes underlying the apprehension and retention of concrete and abstract sentences are qualitatively different. The pattern of errors, both overt and covert, support the supposition that mental imagery plays a role in the internalization of concrete materials but that abstract sentences are mediated by an essentially verbal process. In this, the data are consistent with the findings of earlier studies which have employed selective interference techniques, in particular those of Klee & Eysenck, discussed above. The greater clarity of the present results compared to earlier recall studies appears to be most readily attributable to the greater purity of counting and tracking as measures of verbal and perceptual interference respectively.

The one major finding not accounted for by the present hypothesis was the very poor performance on abstract sentences shown by subjects in both experiments.

This finding has now been reported in a number of studies (e.g. Atwood, 1971; Klee & Eysenck, 1973), and can perhaps be most readily attributed to the initial difficulties in comprehension associated with sentences of this kind (Johnson *et al.* 1972). While it would be unwise to identify the meaning of a concrete sentence with the image it evokes (Brown, 1958), there seems little doubt that the process of understanding is accelerated when imageable referents are present (e.g. Moeser & Bregman, 1972). The differences in memory processing associated with concrete and abstract sentences may be readily conceptualized in terms of the model proposed by Atkinson & Shiffrin (1968). The translation of a sentence into a concrete memory image can be thought of as one 'control process' available to the subject to facilitate the copying of information from the short-term store into the long-term store. This coding process not only facilitates the transfer of information into a more permanent form but also reduces the strain on processing capacity by recoding extended sentences into rich 'chunks' of information, corresponding to mental images (see Begg, 1972). Abstract sentences, on the other hand, cannot be so economically and rapidly encoded. Because of this, each sentence is especially vulnerable to displacement from the limited capacity short-term store, either by the next sentence, or by the backward counting task. The displacement of the verbal code of one concrete sentence by another is not critical for recall as the subject may use an image to reconstruct the original. However, as the meaning of the abstract sentence resides in the words themselves, loss of retention for them leads to the loss of the sentence as a whole. Such an interpretation is consistent with the poor retention reported for abstract sentences and the tendency noted in this study for subjects to transpose nouns from one sentence to another.

The present results also shed light on the locus of the interference effect produced by the interpolated task. Sasson (1971) argued that the effect of his task was to increase the number of response probabilities for the subject at retrieval, in the manner of a classic retroactive interference effect. Atwood, on the other hand, appears to interpret the effect as one upon the registration of the relevant memory code. The fact that the present study has obtained modality interference effects with tasks unrelated to the recalled material appears to militate against any interpretation in terms of response competition. Equally the fact that significant effects are obtained at all positions with sequentially presented triplets of items seems difficult to equate with Atwood's argument regarding interference at registration. Indeed, the present results seem best interpreted in terms of some effect upon the consolidation of the mnemonic information concerned. Just as externally produced icons suffer rapid degeneration, if followed immediately by the onset of a visual masking stimulus (Sperling, 1960), so equally, internally generated images might be vulnerable to incoming perceptual input, assuming that they share the same processing channels (see Brooks, 1968, for evidence on this latter point). Such an interpretation would be consistent with the known effects of subtraction tasks on the recall of verbal material (Reitman, 1971). Clearly, however, while such an hypothesis accounts for the present data reasonably satisfactorily, further research is required, particularly on the effects of verbal and perceptual interference on sentence recognition.

One final comment concerns the differences between sentences in their relative ease of recall. They underline the fact that any theory of sentence memory which is phrased simply in terms of gross differences between verbal and imaginal coding is bound to be an oversimplification. Imagery alone, for instance, would be unlikely to mediate the recall of negative (wolves do *not* follow cattle and sheep) or interrogative (does the attic contain great trunks and also some paintings?) versions of the target sentences, yet changes of this kind are readily detected in recognition tasks (Bregman & Strasberg, 1968). The coding of semantic markers as well as an image seems essential to ensure the preservation of such information. Such parallel coding of verbal and non-verbal information seems to be a common feature of memory tasks involving concrete material (Davies, 1971). It is also conceivable that not all concrete sentences depend to the same degree on imagery. In Expt. II the concrete sentence 'The shoes are a product of the skin of the cow' showed a much lower rate of semantic errors relative to the other two sentences employed. Inspection suggests that this sentence makes a statement about the nature of the world whereas the others convey locational or directive information. In Tulving's terms (Tulving, 1972) the former is representative of semantic memory whereas the latter two sentences are concerned with episodic memory. It seems possible that the role of imagery may be more significant for the latter than the former although all are equally concrete. However, the small size of the sample involved severely limits the scope of such speculations. Further discussion must await the results of a study which systematically varies the nature of the constituent sentences as a function of their semantic and episodic content. This interesting anomaly apart, the present results support the generalization propounded by Begg & Paivio that both verbal and non-verbal processes play a role in the recall of sentences and, further, that such processes may be effectively externalized through the use of selective interference techniques.

REFERENCES

ATKINSON, R. C. & SHIFFRIN, R. M. (1968). Human memory: a proposed system and its control processes. In K. W. Spence & J. T. Spence (eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 2. New York: Academic Press.

ATWOOD, G. (1971). An experimental study of visual imagination and memory. *Cogn. Psychol.* **2**, 290–299.

BANKS, W. P. (1970). Signal detection theory and human memory. *Psychol. Bull.* **74**, 81–99.

BEGG, I. (1972). Recall of meaningful phrases. *J. verb. Learn. verb. Behav.* **11**, 431–439.

BEGG, I. & PAIVIO, A. (1969). Concreteness and imagery in sentence meaning. *J. verb. Learn. verb. Behav.* **8**, 821–827.

BOWER, G. H. (1972). Mental imagery and associative learning. In L. W. Gregg (ed.), *Cognition in Learning and Memory*. New York: Wiley.

BREGMAN, A. S. & STRASBERG, R. (1968). Memory for the syntactic form of sentences. *J. verb. Learn. verb. Behav.* **7**, 396–403.

BROOKS, L. R. (1968). Spatial and verbal components of the act of recall. *Can. J. Psychol.* **22**, 91–96.

BROWN, R. (1958). *Words and Things*. New York: Free Press.

DAVIES, G. M. (1971). The influence of verbal labelling on the retention of picture stimuli: an alternative viewpoint. *J. spec. Educ.* **5**, 323–336.

JOHNSON, M. K., BRANSFORD, J. D., NYBERG, S. E. & CLEARY, J. J. (1972). Comprehension factors in interpreting memory for abstract and concrete sentences. *J. verb. Learn. verb. Behav.* **11**, 451–454.

KIRK, R. E. (1968). *Experimental Design: Procedures for the Behavioral Sciences*. Belmont: Brooks-Cole.

Klee, H. & Eysenck, M. W. (1973). Comprehension of abstract and concrete sentences. *J. verb. Learn. verb. Behav.* **12**, 522–529.

Mehler, J. (1963). Some effects of grammatical transformations on the recall of English sentences. *J. verb. Learn. verb. Behav.* **2**, 346–351.

Meudell, P. (1972). Short-term visual memory: comparative effects of two types of distraction on the recall of visually presented verbal and non-verbal material. *J. exp. Psychol.* **94**, 244–247.

Miller, G. A. (1962). Some psychological studies of grammar. *Am. Psychol.* **17**, 748–762.

Moeser, S. D. & Bregman, A. S. (1972). The role of reference in the acquisition of a miniature artificial language. *J. verb. Learn. verb. Behav.* **11**, 759–769.

Morris, P. E. & Reid, R. L. (1972). Imagery and the recall of adjectives and nouns from meaningful prose. *Psychon. Sci.* **27**, 117–118.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

Paivio, A. & Begg, L. (1971). Imagery and comprehension latencies as a function of sentence concreteness and structure. *Percept. Psychophys.* **10**, 408–412.

Paivio, A. & Csapo, K. (1969). Concrete-image and verbal memory codes. *J. exp. Psychol.* **80**, 279–285.

Paivio, A., Yuille, J. C. & Madigan, S. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *J. exp. Psychol. Monogr. Suppl.* **76** (1, pt. 2).

Reitman, J. S. (1971). Mechanisms of forgetting in short-term memory. *Cogn. Psychol.* **2**, 185–195.

Sasson, R. Y. (1971). Interfering images at sentence retrieval. *J. exp. Psychol.* **89**, 56–62.

Sasson, R. Y. & Fraisse, P. (1972). Images in memory for concrete and abstract sentences. *J. exp. Psychol.* **94**, 149–155.

Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* **74**. (Whole no. 498.)

Sutcliffe, J. P. (1969). The Betts QMI vividness of imagery scale. Cited in A. Richardson, *Mental Imagery*. London: Routledge & Kegan Paul.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (eds.), *Organization of Memory*. New York: Academic Press.

Waugh, N. C. & Norman, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

(*Manuscript received* 13 *January* 1975; *revised manuscript received* 1 *April* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 73–87
Printed in Great Britain

MEMORY DEFICITS IN EARLY INFANTILE AUTISM: SOME SIMILARITIES TO THE AMNESIC SYNDROME

BY JILL BOUCHER
Department of Psychology, University of Birmingham
AND ELIZABETH K. WARRINGTON
National Hospital, Queen Square, London

Autistic children were compared with control children on tasks in which retention was tested by different methods. In three tests of recall, using named pictures, written words and spoken words as test stimuli, autistic children were impaired in comparison with age-matched normal children and with controls matched for verbal and non-verbal ability. In one test of forced-choice recognition of pictures, autistic children were impaired in comparison with ability-matched controls. In three tests of cued recall, using named pictures, written words and spoken words as test stimuli, and acoustic, graphemic and semantic cues, autistic children were not impaired in comparison with normal age-matched controls. In one test of paired-associate learning using non-related word pairs as test stimuli autistic children were not impaired in comparison with normal age-matched controls. These experimental paradigms were similar to some that have been used to investigate the amnesic syndrome in man. Thus findings on paired-associate learning differ in autistic and amnesic subjects, but findings on recall, recognition and cued recall are comparable. A possible parallel between autism and amnesia is discussed.

Autistic children's behaviour is in part defined in terms of repetitive, or perseverative, behaviour: diagnostic criteria include the presence of obsessions, rituals and motor stereotypes (Rutter, 1968). Experimental analysis of autistic behaviour has shown impaired active error reduction and perseveration in learning and recall tasks (Frith, 1970*a*, *b*), reduced exploration (Hutt & Hutt, 1970) combined with increased general activity, and in particular a tendency to initiate and not sustain activities (O'Connor & Hermelin, 1963; O'Connor & Hermelin, 1967), position habits in discrimination learning (O'Connor & Hermelin, 1965) and lack of response to novel stimuli in the presence of familiar stimuli, and reduced spontaneous alternation (Boucher, in press). These observations are analogous to the behavioural deficits reported in animals with lesions of the hippocampus. Hippocampal animals are slow to extinguish (Douglas, 1967) and show motor stereotypes associated with increased general activity (Kimble, 1963). They are impaired in active error reduction (Douglas *et al.*, 1969), and show both reduced exploration and a tendency to initiate rather than maintain activity (Glickman *et al.*, 1970). Spatial reversal learning (Mahut, 1971), response to novel stimuli (Roberts *et al.*, 1962), and spontaneous alternation (Hirsch, 1970) are also impaired.

Lesions of the limbic lobe, including the hippocampus, in human adults may result in the amnesic syndrome, a selective and global deficit of long-term memory, other aspects of cognitive function being intact. It has been suggested that the underlying functional disorder in amnesic patients and animals with hippocampal lesions may be the same, both being accounted for in terms of disinhibition and interference phenomena (Warrington & Weiskrantz, 1970, 1973). The question therefore arises whether there are also similarities between the performance of autistic children and amnesic patients. Such evidence as is available suggests that this may be the case. Autistic children, like amnesic patients, show intact short-term memory (STM) and relatively good performance on tests of visual-spatial reasoning (Bartak *et al.*, 1975). Like amnesic patients they also show relatively good learning and retention of motor skills (Goldstein, 1959; Hermelin & O'Connor, 1970), combined with severe impairment in the learning of other classes of behaviour, notably linguistic and social behaviour. There are, of course, many possible bases for such behaviour patterns: nevertheless, a deficit of long-term memory is one plausible candidate.

Long-term memory functions in autistic children have not been directly investigated. However, experiments on language coding, carried out by Hermelin & O'Connor (1970) and by Frith (1970*a*), used recall tasks with a LTM component, and the autistic subjects showed consistent abnormalities on these tasks. Hermelin (1975) has recently suggested that a basic encoding impairment in STM would cause breakdown of the long-term memory system. Bryson (1970) and DeMyer (1971) have suggested that failures in discrimination learning and visual motor imitation could be interpreted in terms of LTM deficit. A recent experiment by Boucher (1975), showing impaired memory in autistic children for sequences of nonsense pictures, tends to confirm this interpretation.

Evidence against the existence of LTM deficits in autistic children comes from the many anecdotal reports of their remarkably good memory in certain instances (Kanner, 1943; Rimland, 1964). However, Warrington & Weiskrantz's (1970, 1974) recent work with amnesic adults suggests that both grossly impaired and normal memory performance may be demonstrated in these patients, performance being related to the method of testing retention. Reports of good retention by autistic children do not therefore necessarily conflict with the data on amnesic adults. However, if the possible parallel between autism and amnesia is to be pursued, it does become of considerable importance to establish whether or not the pattern of memory performance is qualitatively the same in both diagnostic groups.

The aim of the first two experiments was to establish whether or not autistic children have a long-term memory deficit. Recall and paired-associate learning tests, which are both reliable measures of impaired memory function in older subjects (Welford, 1958) and amnesic subjects (Baddeley & Warrington, 1970; Winocur & Weiskrantz, in press), were given. Three further experiments, which replicate as closely as possible in autistic children Warrington & Weiskrantz's experiments with amnesic patients, compared different methods of testing retention – in these it was demonstrated that retention was impaired in amnesic patients when tested by free recall or by yes/no recognition, but retention was normal when tested by a method of cued recall. If similar differential effects were found to occur in autistic children, the parallel between amnesia and autism would be supported and with it the notion that autistic children might have some basic functional disorder associated with the limbic system.

METHOD

Subjects

Group I: Autistic children. Experimental subjects were all diagnosed autistic on Rutter's (1968) restatement of Kanner's (1943) original definition, and attended schools for autistic children. Of the 29 children tested, 14 took part in only one experiment, six participated in two, three in three, five in four, and one in all five of the experiments.

The sample of autistic children tested in these experiments was of relatively high ability. This selection was made for two reasons. First, if more able children are tested, any deficits are more likely to be attributable to autism than to concomitant mental subnormality. In the second place, it was necessary to select children who could cooperate with formal tests using verbal material. In addition to being of relatively high IQ, only those subjects were selected who were deemed capable of cooperating, and motivated to cooperate, in the experiments. This judgement was made on the basis of performance in easy practice tests and at least some success on some part of the experiment in which the child was participating. This selection of relatively able and cooperative subjects entails that autistic children tested were not representative of the full spectrum of autism and were not necessarily displaying its essential core to the fullest extent.

Group II: Normal children. A total of 28 primary and secondary school children was tested in four experiments using normal control groups. Seventeen of these children took part in only one experiment, six took part in two experiments, five in three, and none in all four experiments. Subjects in the normal control groups were individually paired with the autistic children for sex and chronological age (CA) and were selected by their teachers as being of average ability with no particular problems.

Group III: Ability-matched children. A total of 36 children was tested in four experiments in which ability-matched control groups were used. Eight of this group of children attended normal schools, 21 attended schools for the educationally subnormal (ESN school), and seven attended day special schools for the severely subnormal. Twenty-five of the ability-matched children took part in only one experiment, eight took part in two, three in three, and none took part in all four experiments.

Since autistic children, unlike amnesic adults, are broadly speaking intellectually subnormal, it is important to show that the pattern of memory function, if abnormal, is not merely typical of the subnormal child. Further, since autistic children may show an uneven IQ profile, with low scores on verbal tests and higher scores on non-verbal tests, critical subnormal control groups should be matched with experimental subjects' *lowest* relevant ability (namely language, in these experiments) if relative *impairment* is to be demonstrated in the autistic group, and with the experimental subjects' *highest* relevant ability (namely non-verbal reasoning, in these experiments) if relative *superiority* in the autistic group is to be demonstrated. Matching on both ability criteria simultaneously in experiments in which impairment in one condition and good performance in another are predicted is not always easy, especially if age and sex are also to be matched, as was attempted here. Various compromises between these several constraints, all of which were considered desirable, are represented in the following experiments. Only control subjects who were able and motivated to cooperate were tested.

The Peabody Picture Vocabulary test (Dunn, 1965) was used to match for language level. Austistic children's scores on this test have been found (Bartak *et al.*, 1975) not to differ significantly from scores on the Reynell test (Reynell, 1969) which is a more comprehensive test of receptive and expressive language than the Peabody test. The coloured Progressive Matrices test of non-verbal reasoning ability (Raven, 1963) was used to match for non-verbal ability.

Material and procedure

Test stimuli in the following four experiments were written words, spoken words and pictures. All subjects were given a practice session before each part of each experiment to ensure that tasks were understood. Less able children were given repeated practice on the same material until they gave good evidence of having understood what they were required to do. In those experiments in which the subject repeated each word after the experimenter the test items were presented at a rate of approximately one item per 2 sec. In tasks which required the subject

CALCUTTA

Table 1. *Details of subjects*

	Expt. I		Expt. II			Expt. III			Expt. IV			Expt. V	
	Autistic	Ability-matched	Autistic	Ability-matched	Normal	Autistic	Normal	Ability-matched	Autistic	Normal	Ability-matched	Autistic	Normal
Number ...	17	17	12	12	12	10	10	10	11	11	11	10	10
Sex (M/F)	16/1	16/1	10/2	10/2	10/2	8/2	8/2	8/2	9/2	9/2	8/3	8/2	8/2
Mean CA (yr-mth) and range	9·11 (7·1–14·1)	9·8 (7·3–13·8)	13·7 (8·11–18·4)	13·0 (9·7–16·5)	13 7 (9·7–17·0)	12·9 (8·5–16·11)	12·8 (8·2–16·1)	9·3 (7·1–13·8)	10·9 (7·5–15·4)	10·8 (7·7–15·4)	9·5 (7·2–13·3)	14·2 (9·10–18·0)	13·11 (9·5–16·7)
Mean Verbal (yr-mth) and range	5·5 (3·2–9·2)	Presumed commensurate with non-verbal ability	5·2 (3·3–7·1)	Presumed commensurate with non-verbal ability	Presumed normal for CA	5·5 (3·9–6·3)	Presumed normal for CA	5·7 (3·8–7·6)	5·4 (4·0–7·1)	Presumed normal for CA	5·9 (4·6–7·6)	5·9 (4·7–9·5)	Presumed normal for CA
Mean Matrices,* score and range	22·8 (12–29)	22·3 (15–31)	22·7 (8–33)	22·0 (13–31)	Presumed normal for CA	23·0 (13–32)	Presumed normal for CA	21 3 (19–27)	23·7 (17–32)	Presumed normal for CA	20·0 (15–25)	21·6 (8–34)	Presumed normal for CA

* Scores corresponding to 50 percentile points at different chronological ages are shown below:

CA ...	5½	6	6½	7	7½	8	8½	9	9½	10	10½	11
Score ...	14	15	15	16	17	18	20	22	24	24	26	28

to name a picture or read a word aloud items were presented at a rate of approximately two items per 3 sec. These rates were selected as being the fastest that were practical for all subjects. Retention intervals of 30 sec. or 1 min. were used, these being sufficiently long to eliminate any STM component in the retention task (Glanzer & Cunitz, 1966). The distracting task varied with the ability of the child; normal secondary school children were required to count backwards as fast as possible from a given number; all other children named as fast as possible the numbers or colours (depending on ability) of playing cards dealt out to them. In Expt. III all subjects could read, and the retention interval was filled by reading aloud. All responses were made orally, except in the forced choice recognition test in Expt. I where either a spoken or a pointing response was accepted. In free recall items correctly recalled were recorded. In cued recall, recognition and paired-associate learning, the child's first response was recorded.

Rewards of Smarties for the autistic and younger control children, and halfpennies for the older controls were used. Since reward values across subjects ranging from 17-year-old normals to 8-year-old autistic children could not be equated, no attempt was made to use rewards systematically. All children were praised, encouraged and rewarded for correct responses as much as was necessary to maintain good cooperation (incorrect responses were not rewarded).

EXPERIMENT I

The aim of this experiment was to establish whether or not autistic children have a retention deficit. Autistic children were compared with non-verbal ability-matched controls on a recall task.

Subjects

Seventeen autistic children and 17 children matched with the experimental group for sex, age and non-verbal ability were tested. Six children in the control group attended normal schools, nine attended ESN schools, and two attended a day special school. Further details about the subjects tested in this and the subsequent experiments are given in Table 1.

Material and procedure

The test stimuli were 10 black-and-white outline pictures of common objects. The test stimuli were presented singly at a rate of two items per 3 sec., the subject being required to name each object presented. After a filled retention interval of 1 min., during which children named (according to ability) the colours or numbers of playing cards presented at each subject's maximum naming rate, 2 min. was allowed for recall.

Results

The mean number of items correct and the standard deviations for each group are given in Table 2. The performance of the autistic children on this recall task is significantly inferior to that of the ability-matched controls ($t = 2{\cdot}34$; $P < 0{\cdot}05$).*

Table 2. *Results of Experiment I*

	Autistic children	Ability-matched controls
Free recall		
Mean number of items correct (out of 10)	2·3	3·7
Standard deviation	1·9	1·6

This experiment establishes that retention in autistic children is impaired compared with that of children of comparable age and non-verbal ability when tested by the method of recall. However, since the experimental and control groups were not matched for language ability, it is possible that the poor performance of the autistic children may merely reflect their lower level of language development.

---

* Two-tailed tests of significance are used throughout this paper.

EXPERIMENT II

Autistic children's ability to learn a list of unrelated word pairs over trials was compared with that of a non-verbal ability-matched control group and a normal control group. On the hypothesis of an autism–amnesia parallel, *impaired* learning by the autistic subjects would be predicted (Winocur & Weiskrantz, in press) and the critical control group would therefore be one matched for language ability. However, a pilot study indicated that the prediction might not be fulfilled, and that, on the contrary, *superior* learning by the autistic subjects might occur (Boucher, 1975). The critical control group for demonstrating autistic children's superiority is one matched for non-verbal ability, and such a group was therefore used in this experiment.

Subjects

Three groups of 12 subjects were tested: an autistic group, a control group matched with the autistic children for sex, CA and non-verbal ability, and a normal group. Three children in the ability-matched control group attended normal schools, five attended an ESN school, and four attended day special schools.

Material

Test stimuli were pairs of unrelated words. Each subject was required to learn a list of pairs which was his digit span plus four in length. Factors of meaningfulness, imageability and concreteness (which are known to influence PAL) were not controlled in the material used in this experiment: rather, variety along these parameters was included in order that no single learning strategy (such as the use of visual imagery) could unduly influence the results. For example, the first four word pairs were: chair–lorry; dance–where; fat–blue; much–pen.

Procedure

Word pairs were initially presented orally at a rate of four every 30 sec., the subject repeating each word pair in turn. Stimulus words were then immediately presented again, in the same order as word pairs in the original list, 15 sec. being allowed for recall of any response word. Response errors were corrected as they occurred, the subject repeating the corrected word pairs. No further presentation, after the initial presentation of word pairs, was given, and learning therefore progressed via the correction procedure. A maximum of 10 test trials was given, and the learning criterion was one error-free trial. Those subjects who failed to learn within 10 trials were assigned a trials to learning score by adding 10 to the number of response words not recalled on the subject's best trial. Thus a subject who remembered seven out of 11 responses on his best trial was assigned a score of $10+(11-7) = 14$. The assumption behind this method of assigning scores above 10 was that subjects would have recalled one more word pair on every subsequent trial. Though arbitrary, the assumption allows for different approximations to learning to be reflected in the scores of children who did not reach the learning criterion.

Results

Mean number of trials to learning, and standard deviations, for the three groups of subjects are shown in Table 3. t tests show that autistic children learn significantly faster than the ability-matched control group ($t = 2{\cdot}39$; $P < 0{\cdot}05$) but not significantly faster than the normal children ($t = 1{\cdot}03$). There is no significant difference in the scores of the two control groups ($t = 1{\cdot}59$).

It is striking that not only are the autistic children not impaired relative to the ability-matched controls in this experiment, but they are significantly superior to them, despite being of lower verbal ability. Moreover, the autistic children show as

Table 3. *Results of Experiment II*

	Autistic children	Ability-matched controls	Normal children
Mean Digit Span and range	5 7 (4–8)	4·8 (2–7)	6 2 (5–8)
Mean number of trials to learning	6·0	9·1	7·2
Standard deviation	3·0	3·3	2·6

good ability on the task as normal children of the same age. These findings are in sharp contrast to predictions made on the basis of an amnesia–autism parallel. It might have been expected that the ability-matched control group, which included some severely subnormal children, would have been significantly inferior to the normal children on this task. That they were not suggests that associative learning impairment may be less a feature of subnormality than impairments in other kinds of learning. Large within-group variations are also probably a contributory factor in this lack of significant difference.

EXPERIMENT III

The aim of this experiment was to determine whether the retention impairment demonstrated in autistic children in Expt. I can be accounted for by different language abilities in the autistic and control groups; and to compare different methods of testing retention in autistic children and both normal and ability-matched controls. Two methods of testing retention were used in this experiment: cued recall, using acoustic cues; and forced choice recognition. Forced choice recognition was preferred to a yes/no recognition test (such as had been used in experiments with amnesic patients) since a pilot experiment indicated that autistic children may have difficulty in cooperating with a yes/no test.

Subjects

Three groups of 10 children were tested, autistic children, normal children, and ability-matched controls. In this experiment the ability-matched control group was equated with the autistic children on two criteria, namely language MA, and non-verbal reasoning ability. Nine of the ability-matched control group in this experiment attended an ESN school, and one attended a day special school.

Material

Test stimuli were coloured pictures of common objects. Ten pictures were used for testing retention by the method of cued recall. The names of all the objects in this set of pictures began with a different double consonant, for example FR-uit, CR-own, PL-ate. The double consonant, spoken with a following neutral vowel, (ə), was used to cue recall. All double consonants were common to at least one non-test word of higher frequency in the children's vocabulary; for example, FR-ock, CR-isps, PL-um. This constraint was introduced in an effort to minimize the possibility that correct responses would occur through guessing.

Eighty pictures were used for testing retention by the method of forced-choice recognition: 40 test items and 40 distractor items. All 80 pictures were drawn from the same pool as the items selected for the cued recall condition.

Procedure

All subjects completed the experiment in one test session, cued recall being followed by the recognition condition. In the cued recall condition subjects were shown the 10 pictures singly, naming each picture after the experimenter at a rate of one item per 2 sec. This was done, rather than having the child name the pictures himself, to ensure that the picture name corresponding to the acoustic cue was used at presentation, and also so that the experimenter's voice was heard by the child both at presentation and recall. Retention was tested after a 30 sec. filled interval during which children named the colours and numbers of playing cards presented at each subject's maximum naming rates. The double consonant acoustic cue was presented for each test item in the same order as the pictures were originally presented, 15 sec. being allowed for the recall of each item. In the forced-choice recognition task the 40 test stimuli were presented singly at a rate of two items per 3 sec., the child naming each picture, and retention was tested by the method of forced choice after a 30 sec. filled interval. One test item and one distractor item were presented simultaneously, the subject being required to name or to point to the picture seen previously.

Results

The mean number of items correct and the standard deviations for each group on the recognition and cued recall task are given in Table 4. An analysis of variance was computed; there are significant group differences ($F = 4{\cdot}83$; d.f. $= 2,27$; $P < 0{\cdot}05$), and a significant group × condition interaction term ($F = 7{\cdot}8$; d.f. $= 2,27$; $P < 0{\cdot}01$). On the recognition task the performance of the autistic children is significantly poorer than that of both the ability-matched control group ($t = 2{\cdot}25$; $P < 0{\cdot}05$) and the normal control group ($t = 2{\cdot}89$; $P < 0{\cdot}01$). Ability-matched controls are significantly poorer than the normal children on the recognition task ($t = 3{\cdot}02$; $P < 0{\cdot}01$). Performance on the cued recall task is not significantly different between any of the groups.

Table 4. *Results of Experiment III*

	Autistic children	Normal children	Ability-matched controls
Recognition			
Mean number of items correct (out of 40)	31·0	38·9	37·2
Standard deviation	8·6	0·9	1·5
Cued recall			
Mean number of items correct (out of 10)	7·8	6·9	6 6
Standard deviation	1·5	1·4	1·7

This experiment establishes that there is a retention deficit in the autistic children *as a group* which is not related to poor language ability. Furthermore, differential effects of the method of testing retention in autistic children are demonstrated: forced-choice recognition is impaired in the autistic children compared with both control groups, whereas cued recall is not.

Perhaps the most interesting finding in this experiment, however, is the very large spread of scores which occurs within the autistic group, three children performing at a chance level, four performing as well as normal controls, and the other three autistic children obtaining intermediate scores. This distribution of scores suggests that some, rather than all, autistic children may have recognition difficulties. The large variation in the autistic group technically invalidates the use of parametric statistical tests in analysing results on this experiment. However, non-parametric tests are probably equally misleading in that non-significant findings conceal the fact that half the children in the autistic group performed at a level often well below that of any of the control subjects.

The fact that subnormal children's recognition is inferior to that of the normal control group is not surprising, in view of differences in language development and non-verbal IQ. It is, however, at first sight surprising that subnormals' cued recall is not inferior to that of the normal children. It might also be suggested that the lack of difference between autistic and subnormal children's cued recall scores shows that cueing is not more useful to autistic children than to other subnormals. Two points may be made about relationships between cued recall scores in the three groups. First, it is probable that the subnormal and autistic groups, whose language is less developed than that of the normal children, would use an acoustic coding strategy to a greater extent than normal controls, who would make greater use of semantic coding. Cues that were acoustic would therefore be more relevant in the coding used by autistic and subnormal children than in the coding predominantly used by the normal subjects. This could explain the lack of difference between the normal and subnormal children's cued recall scores. Secondly, the cued recall task in this experiment has a confounding recall component (since some items presented for cued recall would be remembered without cueing), and, if autistic children's free recall is impaired, as is hypothesized, this could explain the lack of difference between the subnormal and autistic children's scores. The implications of these two points are, first that normal children will use semantic cues more efficiently than autistic or subnormal subjects; and secondly that autistic children will use cues more efficiently than subnormals if the confounding recall component is removed. These possibilities are explored in Expt. IV.

EXPERIMENT IV

The aim of this experiment was to compare recall with cued recall using semantic cues, in autistic children, normal children, and a critical language-matched control group, and in an experiment in which the contribution of cueing could be unambiguously assessed. Furthermore, poor recall of meaningful auditory-verbal inputs, combined with good acoustically cued recall, is compatible not only with a theory of amnesic-like retrieval difficulties in autistic children, but also with O'Connor & Hermelin's (1971) theory of impaired coding of meaningful input features. The likely relevance of these two theories to autistic children's cognitive difficulties may be to some extent separated out by testing the ability to use semantic cues: an impaired retrieval hypothesis predicts that semantic cues will be more efficiently used by autistic than subnormal subjects; an impaired input coding hypothesis predicts that the reverse will be the case.

Subjects

Three groups of 11 subjects were tested: autistic children, normal children and children matched with the autistic children for language MA. Ten children in the ability-matched group attended ESN schools and one child attended a day special school.

Material

Test stimuli consisted of three lists of 10 common nouns. Cues were brief semantic descriptions of these words. No cue described more than one test item, but each cue described at least one non-listed word of greater frequency than the test item to which the cue referred. For example, the cue for the test item 'stool' was 'something you sit on', and 'stool' has a lower frequency of occurrence than the non-listed word 'chair' to which the cue is also applicable. This constraint was introduced to minimize successful guessing.

Procedure

The word lists were presented orally at a rate of one item per 2 sec., each word being repeated by the subject. Retention was tested after a 30 sec. filled retention interval during which subjects named the colours or numbers of playing cards. One minute was allowed for recall. Following recall the semantic cues were given for those items *not* already recalled, no specific time limit being imposed.

Results

The mean number of items correct and the standard deviations for each group are given in Table 5. A two-way analysis of variance cannot be applied to these data since the recall and cued recall scores are not independent. Performance in the recall task is significantly impaired in the autistic children compared with both the normal children ($t = 3{\cdot}37$; $P < 0{\cdot}005$) and ability-matched control group ($t = 2{\cdot}17$; $P < 0{\cdot}05$). There is a trend towards significant difference between the normal and ability-matched control groups on the recall task ($t = 1{\cdot}77$; $P < 0{\cdot}10$). However, the total retention score – that is, summing the items recalled and the items correct with cueing – is not significantly different in any of the three group comparisons.

Table 5. *Results of Experiment IV*

	Autistic children	Normal children	Ability-matched controls
Free recall			
Mean number of items correct (out of 3 × 10 = 30)	6·2	12·7	9·6
Standard deviation	4·2	4·9	3·2
Cued recall			
Mean number of items correct (out of 30)	10 2	6·7	7·4
Standard deviation	4·0	1·8	3·4
Total			
Mean number of items correct (out of 30)	16·4	19·5	17·0
Standard deviation	4·6	5·1	4·7

The attempt to assess the relative contribution of recall and cued recall to the overall recall of a single word list raises certain problems. Since cued recall was tested on items not recalled, and recall scores are significantly different in the three subject groups, the number and nature of items subsequently cued differed in the three groups. Thus it is not valid to compare directly the number of cued recall items recalled by each group. However, the finding that the autistic children's total retention scores in this experiment are not significantly different from those of either control group (and there are no ceiling effects), although their performance is worse on the recall task, suggests that autistic children make at least as efficient and possibly more efficient use of semantic cues than a language-matched control group.

This experiment establishes that, in addition to recognition deficits, there are recall deficits in autistic children that are not related to poor language ability. Furthermore, although cued recall scores in the three groups may not be statistically compared, the data suggest that normal children use semantic cues more efficiently than autistic or subnormal children (which would be predicted on the supposition of different coding strategies at different levels of language development), and also that autistic children may use semantic cues more efficiently than subnormals (which would be predicted on an amnesia–autism parallel). The finding that autistic children use semantic cues at least as efficiently as subnormals of comparable language development offers no support for O'Connor & Hermelin's hypothesis of impaired encoding of semantic features.

EXPERIMENT V

In the previous two experiments autistic children's use of acoustic and semantic cues was compared with use of cues by controls, and found to be relatively normal when low language level has been taken into consideration. The aim of the present experiment was to test whether autistic children's good use of cues is modality-specific. In this experiment autistic children's recall and cued recall was compared with that of normal children, using written material and written cues. The use of written material made the selection of an ability-matched (especially of a language-matched) control group impossible, since few subnormal children with poor language development can read, whereas some autistic children do learn to read, though with limited comprehension. Recall was not critically tested in this experiment, therefore, although, in the light of findings on three previous experiments, deficits in the autistic children were expected to occur.

Subjects

Ten autistic children and 10 normal children were tested.

Material

Test stimuli were single words printed on individual cards, words being selected from the Mein (1961) list of the most frequent words in the vocabulary of subnormal children. Three 10-word lists were used in the recall task and three 13-word lists were used in the cued recall task. Cues consisted of the first three letters of each word, written on individual cards. Words for cued recall were all five to seven letters long, and none of them started with the same three letters. All, however, started with the same three letters as at least one other word in the Mein list, and had lower frequency of occurrence than at least one such word. For example, the test item CHA-in has lower frequency than the non-test words CHA-ir, CHA-nge. This constraint was introduced in an effort to minimize the possibility that correct responses would occur through guessing.

Procedure

Recall and cued recall were tested in two sessions on different days, the order in which the different tests were given being alternated between subjects in each group. The word stimuli were presented singly at a rate of two words per three seconds, the subject reading each word aloud. Each list was followed by a filled retention interval of 1 min., during which children read aloud as fast as they were able from *Treasure Island*. In the recall condition, 2 min. was then allowed for recall. In the cued recall condition, cues were presented singly, in a different order to that in which they had been originally shown; no time limitation was imposed.

Results

The mean number of items correct and standard deviations for each group are given in Table 6. An analysis of variance was computed: there is no significant group difference ($F = 1{\cdot}1$; d.f. $= 1, 18$), nor any significant interaction between the groups

Table 6. *Results of Experiment V*

	Autistic children	Normal children
Free recall		
Mean number of items correct per trial (out of 10)	4·0	5·1
Standard deviation	1·8	0·9
Cued recall		
Mean number of items correct per trial (out of 13)	10·7	10·4
Standard deviation	1·2	1·7

and conditions ($F = 2{\cdot}44$; d.f. = 1, 18). However, there is a trend towards significant impairment on the free recall task in the autistic group ($t = 1{\cdot}73$; $P < 0{\cdot}10$).

Normal retention is again demonstrated in the cued recall condition in this experiment, confirming that the efficacy of cues for autistic children is not modality-specific. However, differential effects of the method of testing retention in autistic children are not significant in this experiment, and it is surprising, in view of earlier findings, that the autistic children's recall shows only a trend towards impairment in comparison with normal children of the same age. It seems possible that this was caused by the intervening task being effectively different for the two groups of subjects, and consequent differences in the amount of retroactive interference caused by the task. Normal children may be presumed to code both test stimuli and intervening verbal stimuli predominantly semantically. Autistic children, on the other hand, almost certainly understood little of what they read during the intervening interval, if only for the reason that many of the words were beyond their comprehension. Even if autistic children did code the meaning of test stimuli which were within their comprehension to a lesser extent than the normal subjects, and coded physical features to a greater extent, marked physical differences between the test stimuli and the intervening stimuli would entail that less retroactive interference occurred in the experimental than in the control group. This explanation gains support from the fact that a test of *immediate* recall using the same two groups of subjects and similar test material and procedure did reveal significant impairment in the autistic subjects (Boucher, 1975).

DISCUSSION

This series of experiments has demonstrated that autistic children's memory functions are in some respects qualitatively similar to that of amnesic adults; retention tested by recall and recognition can be impaired, whereas cued recall is consistently normal. These findings hold for pictorial and auditory-verbal inputs, and for acoustic, graphemic and semantic cues.

It is worth noting that the autistic groups were not identical in the three experiments, thus it is unlikely that the positive findings obtained are idiosyncratic to one highly selected sample. Expts. III and IV, comparing recognition and acoustic cueing and recall and semantic cueing, respectively, provide the strongest support for the hypothesis, whereas Expt. V, comparing recall and graphemic cueing, is hardly compelling. However, there are two major findings in this series of experiments that weaken the notion of an autism–amnesia parallel. The first finding concerns the wide spread of scores in the autistic children in tests of recognition and recall. In the recognition test in Expt. III it has been shown that the distribution of scores in the autistic group differs significantly from that in the other two groups. This could suggest that some but not all autistic children may have recognition difficulties. It should be pointed out, however, that ceiling effects occurred in this experiment, and normal recognition in the least affected autistic children cannot therefore be safely inferred. The safer conclusion from this result is that a uniquely wide spread of recognition ability occurs in autistic children which may reflect the underlying heterogeneity of the autistic syndrome.

In some of the experiments with a recall measure (Expts. I and IV) the variance was large in all subject groups. Although the difference in variance between the groups was not significant, it would seem possible that the wide range of age and ability within the control groups was the critical factor, whereas in the autistic group additional factors could be operative. Thus the spread of scores on recall within the autistic group may also, like the spread of scores on the recognition test, at least partly reflect a wide range of specific memory ability in these children. It is worth pointing out that the significant impairment shown in Expts. I, III and IV, despite large within-group variance, strengthens a conclusion of impaired recognition and recall in some, if not in all, of the autistic children.

The second finding in this set of experiments that throws doubt on the validity of an autism–amnesia parallel is the finding of normal learning of unrelated word pairs by autistic children. This conflicts directly with data on amnesic adults, and might be taken to suggest that the hypothesis of qualitative similarity between the memory functions of autistic children and amnesic adults should be rejected.

However, there are reasons for not rejecting the hypothesis out of hand on these grounds. In the first place, although the difference between the performance of amnesic and autistic subjects on this task is a striking one, it could be theoretically explained in terms of expected differences in subjects whose lesions are congenital (or very early acquired) and subjects whose lesions are acquired in adulthood. For the autistic child paired associate learning can easily be accepted as 'new learning', whereas for the amnesic patient the task is perhaps better regarded as one of 'relearning'. That is the autistic child may not have to forget or unlearn old associations which is an essential component of the task for the adult amnesic. Furthermore, there is some recent evidence that amnesic subjects are not impaired on 'new' learning tasks where it can be argued that interference from prior experience is minimal (Baddeley, 1975).

The second reason why the autism–amnesia parallel should not be rejected concerns the positive evidence in its favour which has been reported in this set of experiments, and in the experiment on visual memory (Boucher, 1975). Taking the positive and negative evidence together, the case for regarding the fundamental deficit in at least some autistic children as a congenital or developmental form of the amnesic syndrome must be held as unproven until subjected to further tests. Warrington & Weiskrantz (1970, 1974) have emphasized the differential efficacy of cued recall in amnesic subjects and, to the extent that this phenomenon is obtained in a majority of autistic children, the present hypothesis merits further consideration. The properties and limits of cued recall have been further explored in amnesic patients, and though the basis of the effect cannot claim to have been firmly established, Warrington & Weiskrantz (1974) have argued that the cue operates by restricting the number of available response alternatives in subjects impaired by enhanced persistence/interference of prior learning. Even if qualitative similarities between memory performance in amnesia and autism were substantially confirmed in further experiments, it would be necessary to show that the same causal mechanisms determine the cued recall phenomena in the two classes of subjects. Nevertheless, it is tempting to speculate that a basic deficit in some autistic children may be a memory impairment qualitatively similar to that of the amnesic syndrome in adults. If this were to be established it would provide strong evidence for accepting autism as a physiologically determined syndrome deriving from a congenital defect or developmental anomaly. Furthermore, the implications for the education and management of these children would be far-reaching.

We would like to acknowledge the help of staff and pupils at the Society School for Autistic Children, Ealing; The Helen Allison School, Gravesend; St Nicholas' School and the Castle High School, Kenilworth; St Michael's School and the Ridgeway Day Special School, Warwick; and the Brook Day Special School, Rugby. We are most grateful to Dr R. T. C. Pratt for his comments and advice in the preparation of this manuscript.

REFERENCES

BADDELEY, A. D. & WARRINGTON, E. K. (1970). The distinction between long- and short-term memory. *J. verb. Learn. verb. Behav.* **9**, 176–189.

BADDELEY, A. D. (1975). Theories of amnesia. In R. A. Kennedy & A. L. Wilkes (eds.), *Studies in Long-term Memory*. London: Wiley.

BARTAK, L., RUTTER, M. & COX, A. (1975). A comparative study of infantile autism and specific developmental receptive language disorder. I. The children. *Br. J. Psychiat.* **126**, 127–145.

BOUCHER, J. (1975). A neuropsychological approach to the study of autism; two exploratory hypotheses. (Ph.D. thesis, Birmingham University.)

BOUCHER, J. (In press). Alternation and sequencing behaviour and response to novelty by autistic children. *Neuropsychologia*. (In press.)

BRYSON, C. Q. (1970). Systematic identification of perceptual difficulties in autistic children *Percept. mot. Skills* **31**, 239–246.

DEMYER, M. K. (1971). Perceptual limitations in autistic children and their relation to social and intellectual deficits. In M. Rutter (ed.), *Infantile Autism: Concepts, Characteristics and Treatment*. Edinburgh: Churchill Livingstone.

DOUGLAS, R. J. (1967). The hippocampus and behaviour. *Psychol. Bull.* **67**, 416–442.

DOUGLAS, R. J., BARRETT, T. W., PRIBRAM, K. H. & CERNY, M. C. (1969). Limbic lesions and error reduction. *J. comp. physiol. Psychol.* **68**, 437–441.

DUNN, L. M. (1965). *Peabody Picture Vocabulary Test*. Minnesota: American Guidance Service.

FRITH, U. (1970*a*). Studies in pattern detection in normal and autistic children. I. Immediate recall of auditory sequences. *J. abnorm. Psychol.* **76**, 413–420.

FRITH, U. (1970*b*). Studies in pattern detection in normal and autistic children. II. Reproduction and production of colour sequences. *J. exp. child Psychol.* **10**, 120–135.

GLANZER, M. & CUNITZ, A. R. (1966). Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.* **5**, 351–360.

GLICKMAN, S. E., HIGGINS, T. J. & ISAACSON, R. L. (1970). Some effects of hippocampal lesions on the behaviour of mongolian gerbils. *Physiol. Behav.* **5**, 931–938.

GOLDSTEIN, K. (1959). Abnormal mental conditions in infancy. *J. nerv. ment. Dis.* **128**, 538–557.

HERMELIN, B. (In press). Coding and sense modalities. In L. Wing (ed.), *Early Childhood Autism*, 2nd ed. Oxford: Pergamon Press. (In press.)

HERMELIN, B. & O'CONNOR, N. (1970). *Psychological Experiments with Autistic Children*. Oxford: Pergamon Press.

Hirsch, R. (1970). Lack of variability or perseveration: describing the effect of hippocampal ablation. *Physiol. Behav.* **5**, 1249–1254.

Hutt, C. & Hutt, S. J. (1970). Stereotypes and their relation to arousal: a study of autistic children. In S. J. Hutt & C. Hutt (eds.), *Behaviour Studies in Psychiatry*. Oxford: Pergamon Press.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nerv. Child* **2**, 217–250.

Kimble, D. P. (1963). The effects of bilateral hippocampal lesions in rats. *J. comp. physiol. Psychol.* **56**, 273–283.

Mahut, H. (1971). Spatial and object reversal learning in monkeys with partial temporal lobe ablations. *Neuropsychologia* **9**, 409–424.

Mein, R. (1961). A list of words used in conversation by severely subnormal patients. (Cell Barnes & Harperbury Group Hospital Management Committee.)

O'Connor, N. & Hermelin, B. (1963). Measures of distance and motility in psychotic children and severely subnormal controls. *Br. J. soc. clin. Psychol.* **3**, 29–33.

O'Connor, N. & Hermelin, B. (1965). Sensory dominance in autistic imbecile children and controls. *Archs gen. Psychiat.* **12**, 99–103.

O'Connor, N. & Hermelin, B. (1967). The selective visual attention of psychotic children. *J. child Psychol. Psychiat.* **8**, 167–179.

O'Connor, N. & Hermelin, B. (1971). Cognitive deficits in children. *Br. med. Bull.* **27**, 227–231.

Raven, J. (1963). *The Coloured Progressive Matrices*. London: H. K. Lewis.

Reynell, J. K. (1969). *Reynell Developmental Language Scale*. London: National Foundation for Educational Research.

Rimland, B. (1964). *Infantile Autism: the Syndrome and its Implications for a Neural Theory of Behaviour*. New York: Appleton-Century-Crofts.

Roberts, W. W., Dember, W. N. & Brodwick, M. (1962). Alternation and exploration in rats with hippocampal lesions. *J. comp. physiol. Psychol.* **55**, 695–700.

Rutter, M. (1968). Concepts of autism: a review of research. *J. child Psychol. Psychiat.* **9**, 1–25.

Warrington, E. K. & Weiskrantz, L. (1970). Amnesic syndrome: consolidation or retrieval? *Nature, Lond.* **228**, 628–630.

Warrington, E. K. & Weiskrantz, L. (1973). Short- and long-term memory defects in man. In J. H. Deutsch (ed.), *The Physiological Basis of Memory*. New York and London: Academic Press.

Warrington, E. K. & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia* **12**, 419–428.

Welford, A. T. (1958). *Ageing and Human Skill*. London: Oxford University Press.

Winocur, G. & Weiskrantz, L. (In press). An investigation of paired associate learning in amnesic patients. *Neuropsychologia*. (In press.)

(*Manuscript received* 18 *August* 1974; *revised manuscript received* 28 *February* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 89–100
Printed in Great Britain

CONSTRAINTS ON CLASSIFICATORY SKILLS IN YOUNG CHILDREN

By ROBIN CAMPBELL
University of Stirling

MARGARET DONALDSON
University of Edinburgh

AND BRIAN YOUNG
University of Hong Kong

On three occasions over a one-year period 23 children aged 3–4 were asked to match target objects with other objects which were similar to the target in shape or in colour but not in both. A sequence of matches to each target was asked for. The structure of these sequences is examined for evidence of classificatory abilities. Many subjects appear to be attempting (not always successfully) to follow a strategy of 'holding fast' to their original choice of matching criterion (colour or shape) on subsequent trials. The striking differences between performance after a first-trial colour-match and after a first-trial shape-match are described. It is argued that development during this period involves an increasing ability to ignore the more salient attribute (shape) when matching by colour. The general developmental significance is discussed.

If a person gives evidence of ability to discriminate between objects on some basis of likeness and difference is this enough to warrant the conclusion that he is able to classify? Most writers on the subject – including Inhelder & Piaget (1964), Werner (1948) and Vygotsky (1962), who have contributed perhaps more than any others to our understanding of classificatory skills – would agree that it is not. When a child selects all the red objects then all the white objects and so on from a pile of different colours we can no more be sure that he is classifying than we are justified in ascribing classificatory skill to a bee which systematically discriminates between one type of flower and another. In the definition of classification more than the discrimination of attributes is involved.

Inhelder & Piaget, Werner and Vygotsky have somewhat different ways of stating what more is necessary for true classification and of testing for its presence. Yet they would all agree on this point: that classification is marked at one and the same time by the ability to hold consistently to a criterion of grouping, once chosen, and by the ability to abandon this criterion by a deliberate act of decision and reclassify on a basis that is entirely new. This amounts to saying that, for all three writers, the critical issue is whether the overall behaviour is under the control of an active plan. Is the subject passively noting resemblances or is he actively constructing a classificatory system?

We are concerned in this paper with the development of classificatory skills in children. However, it has often been suggested that there exist parallels between early behaviour patterns and those that are characteristic of pathological mental states in adults. It is therefore interesting to note that Werner (1948) argues strongly for such parallels in the case of classification. He refers to the work of Gelb & Goldstein (1925) who, in speaking of their brain-damaged patients, make very clear the distinction between passive registration and active construction. Gelb & Goldstein say: 'If in sorting they put particular colours together, this procedure is forced upon them passively by the experience of coherence of the moment. Furthermore, even when a particular aspect of the sensory impression has forced this procedure on them, the patients are unable to hold fast to this aspect in their subsequent procedure. They cannot make it into a principle of classification...'

The question we shall now consider is whether our subjects (children between the ages of three and five) are at the mercy of some such momentary 'experience of coherence' when they respond to likenesses and differences in the materials that we present to them, or whether the evidence justifies us in saying that 'principles of classification' already guide the patterns of their responding. It should be noted now that most previous studies of the classificatory skills of children in this age range have tended to the conclusion that true classification is beyond their powers. On this point, too, Inhelder & Piaget, Werner and Vygotsky are at one.

Traditionally, methods of investigating classificatory skills have involved sorting tasks and matching tasks. In the former a collection of objects differing in shape, size, colour, etc., is placed before the subject who is invited to 'put the ones that are the same together' or some such. This is sometimes followed by a second invitation to 'do it another way'. Consistent sorting in both phases of such a task can probably be regarded as a sufficient condition for the existence of classificatory powers. The subject is clearly capable of sorting the material in alternative ways and the consistency of his performance makes it unlikely that his treatment of each object is under the control of whatever property he happens to notice at the time. Unlike Gelb & Goldstein's patients he is able to 'hold fast' to a particular aspect.

Matching tasks constitute a less severe, possibly minimal, test of classificatory abilities. They do not call for the sustained concentration and motor control demanded by sorting tasks and are therefore valuable for studying the early stages of the development of classificatory skills, since pre-school children are often hyperactive, easily distractable, and impatient with lengthy manipulative tasks, the point of which no doubt escapes them.

In matching tasks the subject is required to select an object which matches ('is the same as') some target object. If he is encouraged to pick a number of objects, one after the other, as matches to the same target object in a situation where matches according to more than one criterion are possible, then the nature of the chosen sequence gives evidence about whether he 'holds fast' to one criterion and whether, when the set of objects that satisfy this criterion is exhausted, he is able to recognize that an alternative basis for matching remains.

In this paper we report the results of a longitudinal study using matching tasks as the basis for an assessment of classificatory skill.

METHOD

Materials

Two sorts of stimulus material were used: common objects and formal shapes. The former were pencils, toothbrushes, eggcups and miniature plastic roses, all in a range of colours; the latter were squares, circles, triangles and narrow rectangles, all cut from coloured perspex so that the biggest dimension of each figure was approximately 1 in. The common objects were thus familiar and readily nameable and differed among themselves by many attributes other than the two (shape and colour) which were systematically varied. By contrast the formal shapes – and their names – were likely to be relatively unfamiliar to the children and differed only in shape and colour.

From these materials two sets of objects were constructed as follows:

Task 1: one red, one white and one yellow toothbrush; and similarly for eggcups, pencils and roses.

Task 2: one red, one white and one black square; and similarly for circles, triangles and rectangles.

Thus in each task there were five objects similar in shape or in colour to any given object but no objects similar in both. Two other matching tasks in which the only objects similar to a given object were similar in both shape and colour were also used; but repeated matching was not studied in their case for obvious reasons, and results from them will not be reported in this paper.

Subjects

The subjects were 23 children who were taking part in a larger longitudinal study of the development of cognitive skills. In socio-economic background, the children represented the middle range of the population. They came neither from culturally deprived homes, nor from homes that were specially intellectually stimulating. They were tested on three occasions, at mean ages of 3 years 7 months, 4 years 2 months and 4 years 5 months. The age range of the group was 9 months. The range of IQs on the Stanford–Binet Intelligence Scale before the first test was from 88 to 127 and the mean was 106.

The results obtained from one subject have been discarded because it came to seem probable that he was colour-blind.

Procedure

On each of the three occasions of testing, referred to below as Times 1, 2 and 3, the two matching tasks were administered in the same order: task 1 before task 2.*

On each task, several matching trials were allowed. The procedure for the first trial was as follows. The set of objects was placed haphazardly on a table with the subject facing the experimenter. A target object (task 1 – the red toothbrush; task 2 – the white square) was held up by the experimenter and the subject was asked 'Give me one that's the same as this one in some way'. On second and later trials the subject was asked, 'Now give me another one that's the same as this one in some way'. After each trial the selected object was removed from sight. The target object remained the same throughout the series of trials. Repeated-matching trials ended when the subject failed to find a matching object or when he shifted his criterion. That is, if the first two objects selected in task 1 were toothbrushes, they were presumed to be matches on the basis of shape (referred to below as shape-matches); if now the third object was a red eggcup this was presumed to be a match on the basis of colour (a colour-match). The subject was asked to give a reason for each of his choices.

RESULTS AND DISCUSSION

In the analysis that follows, S stands for a match on the basis of shape, C for a match on the basis of colour and Ø either for choice of an object which did not match (according to our prejudices) or, much more commonly, for a statement that there were no matching objects.

---

* This study was not carried out in isolation but formed part of a much larger one, so some of the features of the present design were determined by the objectives of the wider study. For example, the precise choice of materials was influenced by consideration of certain other tasks which were to be given to the children. And the absence of counterbalancing in the order of task presentation is attributable to the fact that we wanted to be able to make strict comparisons between individual children in the course of longitudinal analysis of the findings.

Table 1. *Sequential matching patterns*

Response pattern		Overall frequency of occurrence
1	Ø	8
2	SØ	7
3	SC	3
4	CØ	9
5	CS	15
6	SSØ	25
7	SSC	13
8	CCS	8
9	CCØ	6
10	CCCS	7
11	CCCØ	31

Table 2. *Distribution of types of response pattern*

	Type of response pattern				
Type of material	No response	P_1	P_2	E_1	E_2
Common objects	5	12	13	28	8
Formal shapes	3	10	13	28	12

Table 1 gives the 11 different patterns of response and their frequencies over all trials and on all occasions of testing. That is, it shows how often, overall, a given sequence of choices occurred. Thus, for instance, the choice of a single object that matched on the basis of shape followed by no further matching choice (pattern 2) occurred seven times.

In the light of the question asked in the introduction, two features of Table 1 are of particular interest: (1) whether the set of objects satisfying the initial matching criterion is exhausted; and (2) whether the criterion ever changes. If the patterns in Table 1 are grouped according to whether or not they possess these two features, then we have the following four types of pattern (excluding pattern 1 in which no matching choice at all is made).

(1) Partial single-criterion matching, in which there is no criterion shift but the single category used is not exhausted. This type is exemplified by patterns 2, 4 and 9 in Table 1. We refer to it in future as type P_1 – that is, partial and involving only one criterion.

(2) Partial double-criterion matching (P_2) in which both criteria are used but neither category is exhausted. This is exemplified by patterns 3, 5 and 8.

(3) Exhaustive single-criterion matching (E_1), in which there is no criterion shift but the single category used *is* exhausted: patterns 6 and 11.

(4) Exhaustive double-criterion matching (E_2) in which the first category used is exhausted and then there is a shift of criterion to the second category: patterns 7 and 10. (It should be noted that we did not persist in asking for more choices after this point to avoid undue lengthening of the task (cf. Procedure), so the reference to exhaustive matching is a reference to exhaustion of the first category only.)

Table 2 gives the frequencies of these four types of pattern and of the residual

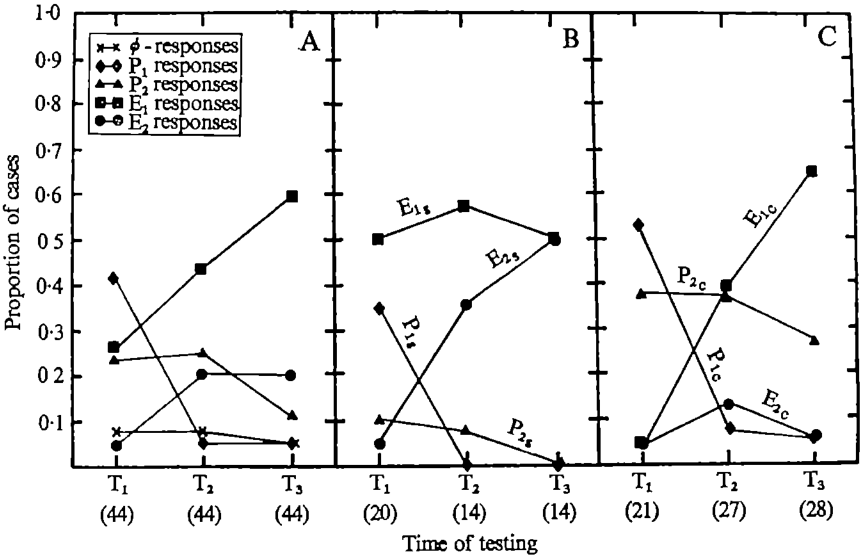


Fig. 1. Proportions of four types of matching pattern at Times 1, 2 and 3.

category of no response for the two kinds of material – common objects and formal shapes.

A χ^2 test reveals no evidence of association between type of material and type of response ($\chi^2 = 1{\cdot}66$; d.f. $= 4$; $P > 0{\cdot}75$). So in the report which follows we present results for the two kinds of material combined.

Single-criterion matching (types P_1 and E_1) occurred in 63 per cent of all cases (excluding, once again, those where no matching choice was made, i.e. pattern 1, Table 1). A notable feature of single-criterion matching is that type E_1 was more common than type P_1 (72 and 28 per cent, respectively) – that is, more often than not, all objects which matched according to the criterion in question were selected. So there is some evidence of a tendency to hold to a criterion once chosen. And it should be noted that there are no grounds for believing that this tendency is due to some enduring preference for shape or colour on the child's part. Colour-matching was slightly more common than shape-matching on initial trials (60 per cent of choices); but at Trial 1 there were always three colour-matches available in the pool and only two shape-matches (see Method). Moreover, our data were not produced by so many individuals who always chose colour and so many who always chose shape. Each child performed this task six times (two types of material by three times of testing) so that there were seven different possible categories of performance (initial match by shape six times, by colour no times; initial match by shape five times, by colour one time, etc.). The distribution of children over these categories is approximately binomial with equal probabilities, as we would expect if the choice of a basis for matching depended on the whim of the moment. Again, those few children who did show an extreme preference for a colour or shape-match on the initial trials were only slightly more prone than the group as a whole to single-criterion matching (68 per cent of cases).

So the 'modal' type of response in our sample is exhaustive single-criterion matching (type E_1) which is characterized by: (1) apparently unsystematic selection of the initial matching criterion; (2) adherence to this criterion once chosen; (3) exhaustion of the relevant category of objects; (4) failure to shift to another matching criterion when the first criterion yields no more choices.

Fig. 1 gives the proportions of the four different types of response at the three times of testing. The suffix 's' appended to a pattern category indicates that the first match was based on shape; and the results for cases where this occurred are shown in Fig. 1B. Similarly the results when the first match was based on colour are shown in Fig. 1C and marked by the suffix 'c'. Fig. 1A gives the two combined. The symbol Ø indicates that no matching choice was made. The numbers below abscissa marks indicate the number of cases on which the proportions are based. Figs. 1B and 1C obviously do not include subjects whose response pattern was pattern 1, i.e. subjects who did not provide either a shape-match or a colour-match at all.

Fig. 1A shows that the frequencies of P_1 and P_2 decline with age while those of E_1 and E_2 increase. Figs. 1B and 1C make it clear, however, that there is considerable interaction between the age trends and the nature of the initial match, for the increase in E_2 is found only when the first match is a shape-match and the increase in E_1 is found only when the first match is a colour-match. Moreover, there is a considerable disparity between the level of P_{2s} responses and that of P_{2c}. These differences might be thought to be explicable on the general ground that more colour choices than shape choices are available and therefore exhaustion of the colour category might be expected to be slower to appear on some assumption that the ability to sustain interest in the task would increase with age. However the disparity between P_{2s} and P_{2c} is maintained over all three occasions of testing so that does not seem to be a particularly helpful explanation. Moreover, it is the level of initial shape responses which is the lower, though initial colour, on average, involves more matches.

Comparison of P_2 with E_1 gives a clue to where a better explanation may be found. When the initial match is a shape-match subjects are more likely to hold to and exhaust a single criterion (E_1) and less likely to shift to the other criterion without exhausting the first category (P_2) than when the initial match is a colour-match. One way to try to account for this is simply to suppose that the children attended more readily to the shape of an object than to its colour, which at first appears perfectly plausible. However, this notion runs into difficulties with two other aspects of the data: (1) the roughly equivalent levels of choice of a first match by colour and by shape (colour being, if anything, slightly more common); and (2) the increase in E_{2s} from Time 1 to Time 3 while E_{2c} remains low throughout. The latter objection is particularly damning since, if shape is the more likely to attract attention, the probability of shifting to shape after exhausting colour should be greater than the probability of shifting to colour after exhausting shape. A somewhat different explanation is therefore proposed.

This alternative explanation is like the first in supposing that shape tends to be in some way a more salient attribute than colour; but it emphasizes not so much attention to a salient attribute as the ability to ignore one. Let us postulate that in order to carry out a successful single-criterion matching process the child has not only to attend to the attribute which forms the basis of the match but also to *ignore* in some positive manner any more salient attribute which does not. The notion is that some attributes are more intrusive than others and must be positively kept out if adherence to a less intrusive one is to be achieved.

This notion is not introduced here as a novel idea. It has a considerable ancestry, though it has not always received the attention it deserves. Külpe (1904) distinguished two perceptual processes of set: positive abstraction in which a crucial aspect of the stimulus is accentuated; and negative abstraction in which the perception of incidental aspects of the stimulus is actively inhibited. What we propose is that processes of negative abstraction are important in the business of 'holding fast' to a criterion *once chosen*, and resisting the temptation to switch to another. It is important to stress that we are not talking here about the initial attractiveness of the different attributes. It is only after an initial choice has been made that it makes sense to speak of the 'intrusion' of some feature of the stimulus – a feature which would have to be ignored if a particular strategy were to be followed consistently.

We propose, then, that the initial choices are largely fortuitous, dependent perhaps on where the child's glance happens to fall as he looks at the array but that differences arise when it is a question of a whole sequence of responses, i.e. of deliberately redirecting attention if the glance happens to fall in the wrong direction.

The following two suggestions seem to fit a number of aspects of our data: first, that the shape of an object is more difficult to ignore in this deliberate and sustained way than its colour;* and second that, with increasing age, subjects become more competent at ignoring.

Then E_{1c} responses will require a deliberate attempt to ignore shape, whereas E_{1s} will call for the less difficult ignoring of colour (which can account for the differences between these two patterns in Fig. 1). If the attempt to ignore shape fails then an (involuntary) shape-match will result (accounting for the $P_{2c}-P_{2s}$ differences). Lastly, even if the attempt to ignore shape is successful, the E_{2c} matching pattern will still be difficult since it calls first for an effort to ignore shape and then for enough flexibility to start attending to it. E_{2s}, on the other hand, looks like a shift in an easier direction.

We may now postulate two ways in which a single-criterion matching process might fail to lead to exhaustion of the category. Either there is failure of some process not specific to the matching task itself (some inadequacy of search of the array perhaps) so that a response of type P_1 results; or else the attempt to ignore one attribute fails, yielding a response of type P_2 – and this is more likely when the attribute to be ignored is shape, with the result that P_{2c} responses are more numerous than P_{2s}. Consideration of the curves for P_1 and P_2 in Fig. 1 makes it seem likely that failures due to the first of these causes are soon eradicated, whereas failures due to the second are still present to some extent even at the third time of testing. These remarks suggest the following developmental pattern.

Type P_{2s} has been omitted from Fig. 2 because the place it should occupy is unclear. In fact, there are only three instances of this response (out of 132 responses) so that, as far as our data can indicate, it matters little where we put it.

---

* Blum & Porter (1973) report work with adult subjects in which active inhibition of attention was triggered by a post-hypnotic cue. They found attention to form harder to inhibit than attention to colour; and they stress the importance of Kulpe's distinction.

$$P_{1s} \longrightarrow E_{1s} \longrightarrow E_{2s} \longrightarrow$$

$$P_{1c} \longrightarrow P_{2c} \longrightarrow E_{1c} \longrightarrow E_{2c} \longrightarrow$$

Fig. 2. Developmental relationships between types of matching pattern. (Cross-sectional analysis.)

We now conclude by examining the longitudinal sequences to see how they relate to the scheme shown in Fig. 2.

The usual technique for investigating longitudinal trends where the data can be ranked is to use Kendall's S as a descriptive statistic and employ, for instance, Jonckheere's test for trend (1967) to test hypotheses of no trend. Since we are dealing with nominal data this cannot be done. Kendall's S compares the rank of items early in a longitudinal sequence with the rank of those that come later, working from early to late through the sequence and summing the differences (positive or negative) in ranks. A comparable approach with nominal data seems to be to take the first item in the sequence and note those which come after it, logging an entry in the appropriate row of an $n \times n$ matrix for each such subsequent item (where n is the number of nominal categories). The next item is then treated in the same way and so on until the end of the sequence is reached. Given a pair of categories, A and B, we can then find in our sample how many times A occurs before B and how many times B occurs before A and compare the two. If this is done for our nine categories $\emptyset$, P_{1s}, P_{1c}, ..., E_{2c} with the Task 1 data and again with the Task 2 data then in many cases the number of before–after comparisons is exceedingly small. If the distribution is 4:0 in favour of our hypothesis, four such comparisons are sufficient to yield a one-tailed P of 0·06 on an exact test of no longitudinal priority, so we will report only those cases where we have four or more such before–after comparisons. With these provisions, we get the results shown in Table 3.

In Table 3, x/y means that of the $x+y$ comparisons, x were 'befores' and y were 'afters'. In brackets, P values on an exact test of no priority are given. These tests are one- or two-tailed depending on the cross-sectional indication. Clearly from the results shown in Table 3 for Tasks 1 and 2 separately, little can be inferred beyond perhaps that P_{1c} and P_{2c} are developmentally prior to E_{1c}. However, the agreement between the results from the two tasks is sufficiently good to justify pooling them. The pooled results in Table 3 yield the information about developmental priority shown in Fig. 3. (In the diagram, an arrow indicates priority.)

If one takes all the priorities suggested by our analysis of the cross-sectional results (cf. Fig. 2) and counts the number of 'befores' consistent with these priorities and the number inconsistent with them, it turns out that 86 per cent of the relevant comparisons are consistent with the cross-sectional analysis, which seems reasonably good agreement.

In an attempt to resolve some of the remaining issues of developmental priority, we now present the data from Tasks 1 and 2, having combined them directly, i.e. before making any analysis of before–after comparisons. This manoeuvre yields for each child a structure consisting of a sequence of three pairs of responses.

From each such structure, three coincidences, 12 'befores' and 12 'afters' are

Table 3. *Before–after comparisons*

	Task 1	
P_{1c} before E_{1c}: 4/0; ($P = 0{\cdot}06$)*	P_{2c} before E_1: 2/4; ($P = 0{\cdot}69$)†	P_{2c} before E_{1c}: 6/0; ($P = 0{\cdot}01$)*
E_{1s} before E_{1c}: 5/1; ($P = 0{\cdot}11$)*	E_{1c} before E_{2c}: 0/4 ($P = 0{\cdot}13$)†	
	Task 2	
P_{1c} before E_{1c}: 4/0; ($P = 0{\cdot}06$)*	P_{2c} before E_{1c}: 9/1; ($P = 0{\cdot}01$)*	E_{1s} before E_{1c}: 4/1 ($P = 0{\cdot}19$)*
	E_{1c} before E_{2s}: 4/1 ($P = 0{\cdot}38$)†	
	Tasks 1 and 2 combined	
P_{1s} before P_{2c}: 4/0 ($P = 0{\cdot}06$)*		
P_{1c} before E_{1c}: 8/0; ($P < 0{\cdot}01$)*	P_{1c} before E_{1s}: 4/0; ($P = 0{\cdot}06$)*	P_{1c} before E_{2s}: 5/0 ($P = 0{\cdot}03$)*
P_{2c} before E_{1s}: 2/5; ($P = 0{\cdot}45$)†	P_{2c} before E_{1c}: 15/1; ($P < 0{\cdot}01$)*	P_{2c} before E_{2s}: 3/0 ($P = 0{\cdot}31$)*
E_{1s} before E_{1c}: 9/2; ($P = 0{\cdot}03$)*	E_{1s} before E_{2s}: 3/1; ($P = 0{\cdot}31$)*	E_{1s} before E_{2c}: 3/1 ($P = 0{\cdot}31$)*
E_{1c} before E_{2s}: 4/5 ($P = 1{\cdot}00$)†		

* One-tailed. † Two-tailed.

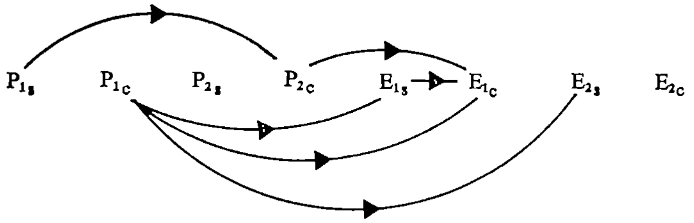


Fig. 3. Developmental relationships between types of matching pattern. (Longitudinal analysis.)

obtained. The results of the analysis will not be presented in full since the matrix is massive and many of the cells still contain small numbers. The omission of such cells yields the half-matrix shown in Table 4. Inside each cell is the distribution of 'befores', coincidences and 'afters' and the goodness of fit χ^2 value for the null expectation of 12:6:12.

It is evident that we have substantial confirmation of the results obtained from the tasks considered singly and we are therefore justified in making further inferences from the pooled data. Also we can, with some confidence, argue that there is no developmental priority of P_{2c} with respect to E_{1s} and of E_{1c} with respect to E_{2s}, an important point in connexion with our presentation of the cross-sectional data relating to these categories given earlier.

The indications of these longitudinal results are summarized in Fig. 4. The horizontal axis represents level of development, the two sequences of categories represent longitudinal priority as assessed by the above analyses, and the bell-shaped curves represent the distribution of cases with respect to level of development and with respect to the response categories (as shown in Fig. 1). It is important to stress

Table 4. *Before–coincident–after comparisons for Tasks* 1 *and* 2 *combined*

	P_{1c}	P_{2c}	E_{1s}	E_{1c}	E_{2s}
P_{1s}	4:2:0 (4·00)	7:0:0 (10 50)	3:1:0 (2·88)	3:0:0 (4·50)	4:0:0 (6·00)
P_{1c}		4:5:2 (4·91)	8:0:0 (12·00)	18:3:2 (14·61)	11:1:0 (13·62)
P_{2c}			5:5:8 (1·31)	28:3:3 (25·63)	5:3:1 (3·22)
E_{1s}				19:1:4 (15·48)	5:2:2 (1·28)
E_{1c}					6:2:10 (2·00)

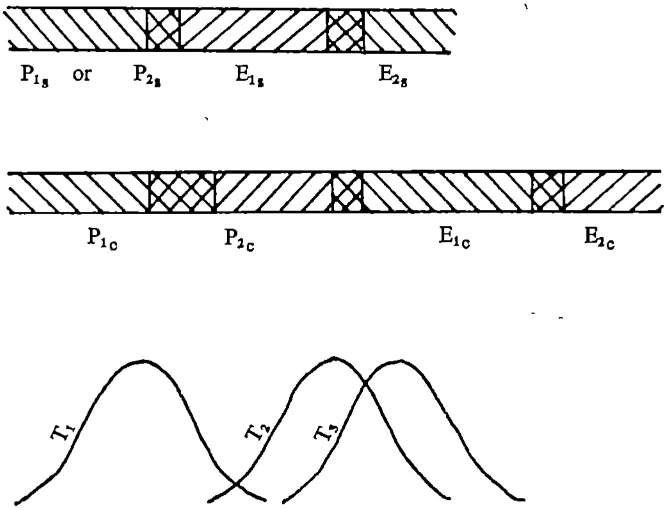


Fig. 4. Schematic model relating level of development to favoured type of matching pattern.

that this is a mere indication – we are unable to find an inferential technique which would allow us to express the degree of support for this model formally. Perhaps the best that can be said for it is that it is maximally consistent with the data obtained.

As a final substantive point about these longitudinal results we would like to draw attention to the fact that, in the model just illustrated, double-criterion matching (P_2 and E_2) occurs at two different levels of development separated by a level at which only single-criterion matching is found. (This kind of developmental pattern is discussed in Donaldson, 1971.) On the other hand, the tendency to exhaust whatever category of objects is selected in Trial 1 of the matching task (E_1 and E_2) becomes steadily stronger as development proceeds.

The main finding, then, is that, even at Time 1, there is a clear tendency for later choices in the sequence to be constrained by the initial one, though the initial choice itself may be largely fortuitous. If a child begins by matching on a given attribute he will tend to hold to that attribute, and, increasingly as he grows older, he will tend to hold to it until no exemplars of it remain.

A fundamental question about the nature of the constraint exercised by the first choice then arises. When a child holds to a single attribute is he following a planned strategy or is he merely 'drawn' to the objects which most resemble the one he has first chosen? Here a first point to note is that the objects which a child selected were removed from his sight before another choice was made. This, of itself, reduces the likelihood that anything like what Gelb & Goldstein (1925) call 'the experience of coherence' could explain our results. However, the fact of central significance is that there tend to be different outcomes according to whether the first choice is a colour-match or a shape-match. The detailed evidence which we have presented suggests that it is necessary to introduce the notion that different attributes have different 'saliency' or 'pull', yet at the same time that the notion of increase in the capacity for voluntary – or planned – resistance to that pull has also to be invoked to explain the developmental data. If a child begins by matching on the basis of shape it is easier for him to hold to this criterion than if he begins by matching on the basis of colour. And this is not so much a matter of being more or less prone to attend to a given attribute as of being more or less able to ignore it.

We have then a developmental lag of the kind shown in Figs. 2, 3 and 4 with the same sequence being followed for shape and colour but with colour showing systematic delay.

To this general statement, however, it is necessary to add a word of qualification. It would be unsafe to conclude that shape is easier to hold to than colour for all shapes and for all colours. There is a growing amount of evidence for the existence of certain kinds of natural perceptual dominance; and among the evidence is a report by Rosch (1973) of findings that there are colours which are more perceptually salient than others; and likewise for forms. If this is so, then our results could perhaps be reversed by constructing the sets of objects from forms of low salience and from colours of high salience. (Both were probably highly salient in the sets we used.)

However this may turn out, if the developmental disparity between shape and colour that we observed is indeed to be explained by postulating that some attributes are harder to ignore than others and that the ability to ignore the more intrusive of them increases over the age-range we studied, then the bulk of our data can be accounted for in terms of successful and unsuccessful attempts to follow a strategy.

Most earlier studies of the classificatory skills of young children have not led to a conclusion of this kind. Notably, the work of Vygotsky (1962) and of Inhelder & Piaget (1964) has been taken to show that children of the age-range we have studied typically do not have a plan that they are attempting to follow. We are therefore in disagreement with them at this point; though, of course, it is crucial not to lose sight of the distinction between matching and sorting.

It looks as if three main types of matching process occur in our sample. In the first of these, there is a single-criterion strategy which, for one or another reason, fails; and this yields a matching pattern of either type P_1 or type P_2. In the second, the single-criterion strategy is successfully carried out; and this is the 'modal' process in our sample, yielding patterns of type E_1. In the third, there is a double-criterion strategy, i.e. the child carries out a single-criterion strategy successfully, then shifts to the other category. This strategy produces response patterns of type E_2, which are rare in our group of subjects unless the first choice is a match on the basis of shape. It must be noted that responses of type P_2 would readily be produced if no strategy of any kind were being followed (and so, of course, fortuitously, would some instances of the other types). No doubt there were a certain number of cases requiring an explanation of this kind. However, once again, the difference between the sequences beginning with shape and those beginning with colour helps one to see what is happening; for if there were no strategy at all, the frequencies of P_{2s} and of P_{2c} would presumably be quite similar. The fact that P_{2s} occurred very rarely (three instances out of 132 responses) argues strongly that, even at time 1 when the children were around 3 years 7 months old, a matching strategy, however unsuccessful, was usually there.

We can therefore conclude that the children's responses are not on the whole like those of Gelb & Goldstein's patients. The procedures they follow are not forced upon them passively. On the contrary they already have some capacity to 'hold fast'; and, at least towards the end of our study, they show signs of an ability to recognize that possible alternative bases for classification can exist. The limitation is that neither of these abilities is as yet free from circumstantial constraints. This is to say that the constraints which the formal classificatory task imposes are still interacting with constraints of a very different kind. The latter – the circumstantial constraints – derive from a kind of force of circumstance that has no relevance for the task. Specifically, they derive in this instance from the child's predispositions ('natural' predispositions perhaps) to respond in various ways to the nature of the material to be classified. The question of whether the matching sequence starts with shape or colour exerts an effect upon the outcome, though it is strictly irrelevant to the requirements of the classificatory endeavour. Until this ceases to be true, it obviously cannot be claimed that the children are, in the fullest sense, capable of classifying.

This study was supported by a grant from the Social Science Research Council. Roger Wales was one of the investigators, and we gratefully acknowledge his contribution to the research.

REFERENCES

BLUM, G. S. & PORTER, M. L. (1973). The capacity for selective concentration on colour versus form of consonants. *Cogn. Psychol.* **5**, 47–70.

DONALDSON, M. (1971). Preconditions of inference. In J. K. Cole (ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.

GELB, A. & GOLDSTEIN, R. (1925). Über Farbennahmenamnesie. *Psychol. Forsch.* **6**, 127–199.

INHELDER, B. & PIAGET, J. (1964). *The Early Growth of Logic in the Child.* London: Routledge & Kegan Paul.

JONCKHEERE, A. R. & BOWER, G. H. (1967). Non-parametric trend tests for learning data. *Br. J. math. statist. Psychol.* **20**, 163–186.

KÜLPE, O. (1904). Versuche über Abstraktion. *Berlin International Congress of Experimental Psychology*, 56–58.

ROSCH, E. H. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350.

WERNER, H. (1948). *The Comparative Psychology of Mental Development.* New York: International Universities Press.

VYGOTSKY, L. S. (1962). *Thought and Language*. Cambridge, Mass: M.I.T. Press.

(*Manuscript received* 25 *April* 1974; *revised manuscript* 4 *February* 1975)

Br. J. Psychol. (1976), 67, 1, pp. 101–112
Printed in Great Britain

THE INFLUENCE OF CONTEXT ON CHILDREN'S USE OF IDENTIFYING EXPRESSIONS AND REFERENCES

By DAVID A. WARDEN

Department of Psychology, University of Strathclyde, Glasgow

Three experiments are reported, each examining young children's ability to use the indefinite article to introduce a new referent to a context of discourse, and the definite article to refer to an already-introduced referent. The context of the speaker's referring expressions is varied in terms of the nature of, and the listener's knowledge of, the referents. The results indicate that children under five years fail to take account of their audience's knowledge of a referent – their referring expressions are predominantly definite. Between five and nine years, children inconsistently introduce referents with indefinite expressions. It is argued that mastery of the introductory function of *a* requires decentration on the part of the speaker.

In any study of children's language, it is necessary to examine not only the syntactic and semantic aspects of sentence structure, but also the relationship between a verbal utterance and its non-verbal context. It is obvious, for example, that non-linguistic variables such as the speaker's intention and the listener's expectation exercise considerable constraints on successful communication. These contextual variables are particularly relevant when one examines the use of referential language. From a psychological point of view, referring is analogous to pointing, insofar as the referential capacity of language enables a speaker to point to an item verbally by means of a *referring expression*, e.g. 'That man...'; once an item has been picked out in this way, a speaker may comment upon it by using a *predicative expression*, e.g. '...bought my car'.

Many writers (e.g. Gruber, 1967; Bloom, 1970) have noted that very young children, under 30 months, seldom use referring expressions, particularly when the referent is manifest; rather their utterances are predominantly predicative. Moreover, when a child begins to refer, his referring expressions often fail to identify his referents for a listener; for example, children use pronouns such as *he* and *it* in a notoriously careless manner.

In order to produce an appropriate referring expression, which a listener will understand, a speaker must select his expression from a number of alternatives so that it is appropriate to the context of the discourse. He must decide, for example, whether his referent is something that requires introduction and explanation, or whether it may be assumed to constitute common knowledge between himself and his audience. His choice of referring expression will also be influenced by his intention; for example, he may wish to emphasize his referent (*this* book), or indicate to whom it belongs (*my* hat), or he may merely wish to name the referent (*a* pterodactyl). A child must therefore master not only the variety of resources for making a reference in English, but also the cognitive factors that determine which referring expression is appropriate in a particular context. The experiments reported here represent an attempt to study one aspect of children's use of referring expressions, namely their use of the definite and the indefinite article.

It has been traditionally supposed that the articles *the* and *a* represent the conceptual opposition between definiteness and indefiniteness. However, this contrast forms only part of their function, and a fuller linguistic analysis reveals several semantic uses and interpretations (e.g. Christophersen, 1939; Strang, 1962; Warden, 1974). For example, the indefinite article may be used not only in an indefinite sense ('I need a hammer'), but also in a nominative sense ('That is a hammer') and an identifying sense ('There is a hammer in that drawer'). In the first case, the reference is to any member of the class of hammers; in the second case, a particular item is described as being one member of the class of hammers; and in the third case, a particular hammer is being introduced to a context of discourse for the first time. Although the nominative and identifying uses of *a* are similar, in that both are concerned with the marking of unity rather than the marking of indefiniteness (Perlmutter, 1970), naming an item and identifying it remain distinctly different operations. For example, in the nominative sentence 'That is a hammer', the indefinite noun phrase does not identify anything – the referent has already been identified and is now referred to by the expression 'That'; whereas in the sentence 'There is a hammer in that drawer', the indefinite noun phrase identifies a particular item. Furthermore, the nominative sense of *a* is not affected by the contextual factors which constrain the use of the identifying sense of the article; a speaker need not take account of his listener's prior knowledge of a referent when he is naming it; whereas this is precisely what he must do if he wishes to refer to it. When a speaker wishes to refer to a previously unidentified referent, he must first introduce it to the discourse by using an indefinite noun phrase, or identifying expression (Vendler, 1967); once an item has been identified in this way, it may be referred to with a definite noun phrase. It is with this complementary use of the articles that the following experiments are concerned, and the terms *identifying expression* and *reference* will henceforth be used to indicate these respective uses of indefinite and definite noun phrases. The first experiment also compares children's ability to use *a* nominatively and in identifying expressions.

The age range of the children in these studies was partially dictated by the feasibility of conducting controlled experiments with very young children. Naturalistic evidence (e.g. Bloom, 1970; Brown, 1973) suggests that function words such as the articles appear in children's spontaneous speech around three years of age. In an earlier study (Warden, 1974), an attempt was made to elicit referential speech from children between three and five years old, in a controlled experimental setting. However, it was clear that relatively few children below four years were capable of directing their use of language to such non-egocentric goals as describing events to an audience. In the first two experiments reported, therefore, four-year-old children were used as subjects.

EXPERIMENT I

The aims of the first experiment were threefold. First, the experiment was designed to discover whether young children are sensitive to the identifying function of the indefinite article, and to the anaphoric function of the definite article, namely, referring back to a previously identified item. Do young children identify specific referents with an indefinite noun phrase, or is the process of 'pointing' too intrinsically definite for them to realize the need for an indefinite expression? Second, it was necessary to establish that adults, whose language is the model towards which the child is developing, do, in fact, make appropriate use of identifying expressions and definite references. And third, children's ability to use the indefinite article nominatively and in identifying expressions was compared. Naturalistic evidence (e.g. Bloom, 1970) suggests that children first begin to use the indefinite article when they are naming an item. It is predicted, therefore, that children will master the nominative use of *a* before they use this morpheme appropriately in identifying expressions.

Method

Twenty children, between the ages of 4 years 3 months and 4 years 9 months, and 20 adults, average age 20 years, acted as subjects. Only native speakers of English were used.

A farmyard scene was arranged on a table, using the following sets of identical model animals: three horses, four cows, four pigs, four hens, four ducks and four sheep. A portable tape-recorder was concealed under the table to record the subjects' verbalizations.

Each subject was presented with four tasks. Two of these tasks (the description tasks) required him to describe an action involving two model animals, manipulated by the experimenter, in which one animal knocked the other animal down: the other two tasks (the naming tasks) required the subjects to name a previously unidentified animal indicated by the experimenter. The order of presentation of these four tasks was counterbalanced using a simple ABBA design, with one half of the subjects receiving a description task first (Group 1), and the other subjects (Group 2) receiving a naming task first. As the purpose of the experiment was to discover whether the subjects could use identifying expressions appropriately in the description tasks, a stratagem was necessary to ensure that such expressions were required in the experimental context. When presenting these tasks to the children, therefore, the experimenter blindfolded himself and manipulated the animals from memory. The children were instructed as follows: 'I'm going to make some of the animals move about and do things, and I want you to tell me what's happening. I'm going to put this mask on, so I won't see anything, and I won't know what animals are moving or what they are doing. I want you to tell me.' It was hoped that this strategy would also induce the children to describe the events fully, and not merely point to the referents. The blindfold was not used with adult subjects, who were merely asked to describe the events, but to imagine an audience who could not see them. In the naming tasks, the experimenter picked up an animal and asked 'What's that?'

To perform appropriately on the four tasks, the subjects should use the indefinite article in all their referring expressions. In the naming tasks, they should respond to the question 'What's that?' with a reply of the form '*a+noun*'; and in the description tasks, their responses to the instruction 'Tell me what's happening' should be of the form 'A+noun is ——ing a+noun' (e.g. 'a cow is chasing a duck').

Results

The subjects' referring expressions were scored as either definite or indefinite: definite noun phrases usually contained the definite article, although the demonstrative article *that* was used occasionally, and the pronoun *he* appeared three times in the children's descriptions: indefinite noun phrases contained the indefinite article, although the numeral *one* was sometimes used in the description tasks. The proportional use of indefinite expressions by adults and children in the two tasks is shown in Tables 1 and 2. In both these tables, Group 1 subjects received a description task first, and Group 2 subjects received a naming task first.

Clearly, there was no age difference in the use of articles in the naming tasks; both adults and children consistently used the indefinite article (see Table 1). Comparing the children's use of the indefinite article in the naming and description tasks, the

Table 1. *Proportional use of the indefinite article in naming tasks*

	Group 1	Group 2
Adults	1·00	1·00
Children	1·00	1·00

Table 2. *Proportional use of identifying expressions in description tasks*

	Group 1	Group 2
Adults	0·50	0·80
Children	0·12	0·30

prediction that children would master the nominative use of *a* before its use in identifying expressions would seem to be borne out; using the Wilcoxon matched-pairs signed-ranks test, it was found that the children used the indefinite article significantly more often in the naming tasks than in the description tasks ($n = 18$; $T = 0$; $P < 0{\cdot}005$, one-tailed). Their descriptions were marked by the use of inappropriate definite references. The adults, as predicted, used identifying expressions more often than the children (see Table 2). By scoring the number of identifying expressions used by each subject, and applying a Kendall's S test for a condition of one dichotomy, this age difference proved to be significant ($S = 240$; $P < 0{\cdot}001$, one-tailed).

One unpredicted result may best be described as a 'task order effect'. Apparently, the order of presentation of the four tasks affected the use of identifying expressions. Group 2 subjects, who received a naming task first, were more likely to use an identifying expression in the description tasks than Group 1 subjects, who began with a description task (see Table 2). Although this group difference was not significant, the fact that it was apparent in the descriptions of both the adults and the children suggests that it was not just a chance result. It would appear that the act of naming created a set for the indefinite article, thereby increasing the likelihood of an identifying expression in subsequent discourse.

Finally, we had expected the adults to use identifying expressions consistently in the description tasks. The fact that they did not do so suggests that the contextual conditions under which a speaker realizes the need for an identifying expression were not met in this experiment. A speaker need not identify an item verbally if his audience can see it; he may refer to it immediately. In this case, although every effort was made to impress on them that they were to describe the events to an audience who could not see the referents, the adults may have assumed that the referents were already identified, that they were talking to the experimenter, who was moving the animals anyway. The children, too, may have assumed that the experimenter, even blindfolded, somehow shared their view of the events. If so, this assumption could have affected their use of the articles, although children's use of referential language may be less sensitive to such contextual factors. The following experiment was intended to obviate this weakness of design.

EXPERIMENT II

The main aim of this experiment was to discover the extent to which adults and four-year-old children take account of the social context when they are constructing a referring expression. A definite reference to a previously unidentified item is acceptable if the speaker and his audience are looking at the referent together; but when a speaker mentions a previously unidentified item which his audience cannot see, he must use an identifying expression. It is predicted that adults are sensitive to the social context and will vary their use of identifying expressions and definite references appropriately.

In the previous experiment, the children's referring expressions were marked by an absence of identifying expressions and an inappropriate use of definite references. We now wish to discover whether or not children's referring expressions are sensitive to changes in the social context. In other words, did the results of Expt. I reflect the children's assumption that, in that context, the referents were already identified, or do children consistently fail to use identifying expressions, regardless of context?

This experiment also examined the hypothesis that a previously unidentified referent is more likely to be referred to if it is isolated rather than if it is a member of a group. It is suggested that the group situation is more likely to emphasize the need for an identifying expression when talking about one member of that group.

Method

Thirty-two adults and thirty-two children, all native speakers of English, comprised the two experimental groups. The children were in the age range 4·0–4·8; the adults were again students, average age 20.

The experimental stimuli were sixteen line drawings of animals, drawn in black ink on white card. Each drawing depicted a scene in which one animal chased another animal across the picture. Four different pairs of animals were used: a cat chasing a mouse; a horse chasing a cow; a duck chasing a sheep; and a dog chasing a hen. Each of these animal pairs appeared on four drawings, each time in a different setting. For example, there were four drawings in which a cat chased a mouse; in one of these drawings, three cats were sitting in the background of the picture; in another, three mice were shown in the background; while in a third, three mice and three cats were shown sitting in the background; only in the fourth drawing were the cat (agent) and mouse (object) the only animals present.

The sixteen drawings were divided into four groups of four such that each animal pair appeared once in each group. Also, each group contained one drawing depicting each of the four settings. For example, the four drawings in one group were:

1. A cat chases a mouse.
2. A horse chases a cow (with three cows in the background).
3. A dog chases a hen (with three dogs in the background).
4. A duck chases a sheep (with three ducks and three sheep in the background).

In addition to these 16 drawings, one further picture was used as a trial card. This card showed four foxes all chasing four rabbits.

Each subject received one group of drawings, presented one at a time. One half of the subjects (16 adults and 16 children) looked at the drawings with the experimenter (the Social condition), and the other subjects looked at the drawings by themselves (the Isolated condition). The presentation of the four groups of drawings was counterbalanced over subjects, as was the order of presentation of the drawings within each group.

The trial card was shown first, and each subject was asked to describe what was happening in the picture. During this trial, the subject and the experimenter looked at the drawing together.

Table 3. *Mean number of definite referring expressions*

	Isolated condition	Social condition
Adults	2·1	3·4
Children	4·4	5·0

Table 4. *Mean number of indefinite referring expressions*

	Isolated condition	Social condition
Adults	3·3	3·4
Children	2·1	2·1

Table 5. *Mean number of undetermined referring expressions*

	Isolated condition	Social condition
Adults	2·6	1·2
Children	1·5	0·9

One of the four groups of drawings was then placed face downwards in front of the subject, who was given the instructions appropriate to his experimental condition. In the Social condition, subjects were told, 'Now we're going to look at some more pictures, and you must tell me what is happening in the pictures'; in the Isolated condition, subjects were told, 'Over there are some more pictures; I want you to go over and pick them up one at a time, look at them by yourself, and then tell me what is happening in the pictures.' The subjects' descriptions were recorded.

Results

Referring expressions were classified as either definite, indefinite or undetermined. Definite noun phrases usually contained the definite article, although the children also used the demonstrative articles *that* and *these*, and the pronouns *him*, *it*, *they* and *them*. Indefinite noun phrases included the determiners *a*, *one* and *another*. Although undetermined referring expressions are usually classified as indefinite, the frequency of occurrence of such expressions, e.g. '*Cat* is chasing *mouse*', made it necessary to include a separate 'undetermined' category. Tables 3–5 show the mean frequencies of each of these three categories of referring expressions, as they were used by adults and children in both experimental conditions. The data in these tables are not, of course, independent: each subject contributed eight referring expressions, and so the scores in any one cell sum to eight over the three tables.

The adults used fewer definite and more indefinite referring expressions than the children, in both experimental conditions (see Tables 3 and 4). Scoring each subject for the number of definite and the number of indefinite referring expressions he produced, and applying Kendall's S test for a condition of one dichotomy, four comparisons were made. The adults and children in the Isolated condition were compared for their use of definite referring expressions and indefinite referring expressions, and similar comparisons were made for the two groups in the Social condition. Only one of these group comparisons yielded a significant result: in the Isolated condition, the adults produced significantly fewer definite references than the children ($S = 123$; $P = 0{\cdot}007$, one-tailed). The absence of a group difference in the use of definite references in the Social condition does accord with our hypotheses. The adults used definite references considerably, though not significantly, more often in the Social condition than in the Isolated condition (Kendall's $S_o = 57{\cdot}7$; $P = 0{\cdot}108$, one-tailed); whereas the children produced approximately the same frequency of definite references in both conditions (Kendall's $S_c = 6{\cdot}7$; $P = 0{\cdot}45$, one-tailed). In other words, the adults responded to variation in the social context more than the children did, at least as far as their use of definite references was concerned. However, the different experimental contexts did not influence the frequency of identifying expressions produced by either group. This was surprising in the case of the adults, who were expected to produce such expressions more frequently in the Isolated condition than in the Social condition. Concomitant with the relatively low frequency of indefinite articles produced by adults in the Isolated condition (see Table 4) was a surprisingly high frequency of undetermined referring expressions produced by the same group (see Table 5). These undetermined expressions were invariably subject noun phrases, and the omitted determiner was therefore the first word in the sentence.

There was no significant difference between conditions in the children's use of either definite or indefinite referring expressions. However, one difference did emerge from their descriptions: a significantly higher frequency of context-bound, deictic (i.e. pointing) words occurred in the Social condition. Such words as *that*, *him*, *it*, *they*, which were classified as definite referring expressions, were used sixteen times by children in the Social condition, and only three times by children in the Isolated condition. The adults did not use context-bound referring expressions in either condition. Scoring each child for the number of such expressions that he used, and applying Kendall's S test for a condition of one dichotomy, this condition difference was significant ($S_c = 96$; $P = 0{\cdot}028$, two-tailed). As there were no other reliable differences in the children's descriptions, the scores in the two conditions were combined, and a Wilcoxon matched-pairs signed-ranks test showed that children used definite referring expressions significantly more often than indefinite referring expressions overall ($n = 27$; $T = 69{\cdot}5$; $P < 0{\cdot}002$, one-tailed).

It was suggested that definite references might be more likely to occur if the referent was isolated, as the presence of other group members might emphasize the need for an identifying expression. The results disprove this hypothesis; neither the adults nor the children discriminated reliably between isolated and non-isolated referents in their use of definite and indefinite referring expressions.

Finally, the adults were expected to use some definite references in the Social condition, and this they did. But they also used definite references in the Isolated condition, when they were expected to use only identifying expressions. Once again, the adults seem to have made certain presuppositions regarding their audience's knowledge of the referents, presuppositions which the experimental task was designed to eliminate. In order to discover whether children can use identifying expressions appropriately, it is necessary to design a task in which adults consistently use identifying expressions. In an attempt to achieve this, the experimenter took no part in the communication process in the next experiment; experimental subjects acted as speaker and listener.

EXPERIMENT III

This experiment was designed to eliminate such contextual variables as 'shared perception' and 'prior knowledge' from the referential context. Each subject was presented with a cartoon story which he was required to tell to another subject who could not see the drawings. By using a story-telling situation, it was hoped to remove the repetitive element which was present in Expt. II, and which may have induced a stereotyped, and unnatural, use of referring expressions. It was also decided to study a wider age range of children, and examine developmental changes in the use of definite and indefinite referring expressions. It was predicted that, as a child gets older, he will use progressively fewer inappropriate definite references to introduce a new referent.

Method

Four groups of children, in the age ranges 3·0–3·11, 5·0–5·11, 7·0–7·11, 9·0–9·11 respectively, and one group of adults, average age 20 years, acted as subjects in the experiment. There were 16 subjects in the three-year-old group, 20 subjects in each of the five-, seven- and nine-year-old groups, and 10 subjects in the adult group.

Two cartoon stories, drawn in black pencil on white cards, were used. Each story comprised three sequential events, and each event was drawn on a separate card. The two stories were matched in referential content: each contained three animate referents and one inanimate referent; each referent appeared in at least one picture, and at least two referents reappeared in a second picture. It was intended that each subject would mention four referents once, and at least two referents a second time, in the course of telling the story. The two stories may be described as follows:

Story A: A dog is chasing a hen (picture 1). A cow stops the dog, and the hen is hiding behind the cow (picture 2). The hen has laid an egg (picture 3).

Story B: A cat is walking under a tree, and a bird is sitting in the tree (picture 1). A dog chases the cat up the tree (picture 2). The bird is flying away (picture 3).

The subjects were tested in pairs, with the subjects in each pair being selected from the same age group. Subjects were seated at opposite ends of a table, and a screen was placed between them to prevent the listener from seeing the cartoons, and to emphasize to the speaker (storyteller) that his audience was ignorant of the content of his story. The two subjects took turns as speaker and listener; one subject told one of the cartoons to his partner, and then the roles were reversed, and the other subject told the second story. The order of presentation of the two stories was alternated. The three cards of each story were placed, face down in the correct order, in front of the storyteller, who was instructed to turn them over and tell the story to his partner. The listener was instructed to attend carefully to the story, and try to remember it, because he was going to be asked to repeat it afterwards. Although listeners were asked to repeat the stories they heard, in order to maintain face validity, the purpose of this instruction was to emphasize to both subjects, and in particular to the speaker, that they were to tell the stories to each other. The subjects' stories were tape-recorded for subsequent transcription.

Results

Almost every subject mentioned the four referents in his story once, and at least two of the referents a second time. The subjects' referring expressions were scored as definite or indefinite, and as a first or a second mention of the referent. Most definite and indefinite expressions contained the definite and the indefinite articles respectively; undetermined noun phrases were scored as indefinite, and context-bound references (e.g. *he, it*) were scored as definite.

There was no reliable difference in the pattern of responses to stories A and B,

Table 6. *Definite and indefinite referring expressions in the story-telling task*

Subject group	First mention Definite expression	 Indefinite expression	Second mention Definite expression	 Indefinite expression
Adults	—	100	100	—
Nine-year-olds	18	82	100	—
Seven-year-olds	39	61	100	—
Five-year-olds	38	62	90	10
Three-year-olds	54	46	92	8

and the scores from these stories were combined. These scores, expressed as percentages of the total number of referring expressions produced, are shown in Table 6. Table 6 shows that, for example, when the nine-year-olds mentioned a referent for the first time, 82 per cent of their referring expressions were indefinite; this percentage represents 66 out of a total of 80 referring expressions produced by that group (20 subjects × 4 referents).

Two results stand out in Table 6. First, the adults always used identifying expressions and references appropriately. Second, there were very few age differences in the type of referring expressions used to mention a referent for the second time; almost all subjects, from three years upwards, used a definite reference. It was when a referent was being mentioned for the first time that subjects' referring expressions differed, and consequently statistical analysis is directed to these differences.

Between group differences. Each subject was scored for the number of indefinite referring expressions he used to introduce a referent. A Kruskal–Wallis one-way analysis of variance revealed significant differences between subjects on this measure ($H = 48{\cdot}9$; $P = 0{\cdot}001$). The Mann–Whitney U test was applied to each adjacent pair of subject groups in order to identify the source of this significant difference; in fact, every group difference except that between the five- and the seven-year-olds was significant in the expected direction. Adults used more indefinite expressions than nine-year-olds ($U = 45$; $P < 0{\cdot}01$); nine-year-olds used more than seven-year-olds ($U = 111$; $P < 0{\cdot}01$); seven-year-olds used more than three-year-olds ($U = 119$; $P < 0{\cdot}025$); and five-year-olds used more than three-year-olds ($U = 106$; $P < 0{\cdot}01$).

Within group differences. The sign test and the Wilcoxon matched-pairs signed-ranks test were used to compare the frequency with which definite and indefinite referring expressions were used by the subjects in each group to introduce a referent. The three-year-olds were the only group to use more definite than indefinite referring expressions at first mention, but this difference was not significant. All the other groups made greater use of indefinite expressions. This difference was significant for the adults ($n = 10$; $P < 0{\cdot}001$, one-tailed) and the nine-year-olds ($n = 18$; $P < 0{\cdot}0001$, two-tailed), but it was not significant in the seven- or the five-year groups.

A number of qualitative aspects of the children's stories are noteworthy. Table 6 shows that the three- and five-year-olds used a fairly high proportion of indefinite expressions at first mention. However, 70 per cent of the three-year-olds' indefinite expressions and 45 per cent of those produced by the five-year-olds occurred in nominative sentences, e.g. 'It's a cat', and do not therefore constitute identifying expressions, because naming an item presupposes its identification. Another point of interest lies in children's inconsistent use of the articles, particularly in the five- and seven-year groups. Several children appeared to understand the rule which requires a referent to be identified with an indefinite expression, and introduced a referent in this manner; but they then immediately introduced another referent with a definite reference, e.g. 'A cat jumped up a tree and the dog came along'. Furthermore, in the same age groups, namely five and seven years, several children made spontaneous self-corrections in their use of articles; each correction was of the same type, namely, the child would introduce a referent with a definite reference, hesitate, and then repeat his referring expression using the indefinite article. It was also observed that one highly sophisticated three-year-old used identifying expressions and references appropriately throughout his story.

GENERAL DISCUSSION

The adults only used the articles consistently correctly in Expt. III and therefore it may be inferred that certain contextual assumptions did affect their referring expressions in Expts. I and II. As these factors may also have affected the children's referring expressions, the results of Expt. III should offer the best indication of children's referential ability. In fact, these results generally supported the results of the previous experiments; younger children used significantly more inappropriate definite references than older children. What conclusions can therefore be drawn from these experiments, regarding young children's ability to use the articles appropriately? In the first place, children master the nominative use of *a* before its use in identifying expressions. When they named an item, four-year-old children used an indefinite expression (Expt. I), but when they placed a referent in a particular context, by describing an event, they revealed a significant tendency to use a definite reference (Expts. I and II). The children seemed to be operating with an article usage rule stating that referents are specified by virtue of their participation in an event, for example, 'The cow is running' – Which cow? – 'The one that is running'. Young children fail to take account of the social context of their reference, or of their audience's knowledge of the referent, when they construct a referring expression. They fail to recognize the need for an indefinite expression when introducing a referent for the first time in a discourse; consequently, they also fail to recognize the constraints on the use of the definite article, namely that its use indicates an already-identified referent.

The most obvious explanation for a child's failure to identify referents is that he is unable to adopt his audience's point of view. From his own egocentric viewpoint, a referent is specified as soon as he (the speaker) is familiar with it; he fails to realize that his audience will only become familiar with his referent after he has identified it for them verbally. And the results of Expt. III suggest that a child reaches nine years of age before he reliably identifies his referents for an audience. However, if children's failure to identify referents is attributed to their egocentricity, why did nearly every child from four years upwards produce at least some identifying expressions? Were they only partially egocentric? Furthermore, why did the children who used context-bound referring expressions in Expt. II only do so in appropriate contexts, namely, when their audience could see the referents? It may be argued that five-year-old children can be non-egocentric in their use of referring expressions, for example, when using demonstratives; but that they are still grappling with the implications of non-egocentricity for the use of the articles. It seems likely that children's difficulty with the articles stems from the dual function of the indefinite article, namely, to indicate either an indefinite referent or a specific, but previously unidentified, referent. In the former case, a speaker need only consult his own knowledge of a referent, whereas in the latter case he must take account of his listener's knowledge. Children may be forced to rely on the definite article until they have mastered the identifying function of the indefinite article; and this mastery will depend on an awareness of their audience's point of view. The fact that most children sometimes used the indefinite article in their descriptions may indicate their partial awareness of the need to identify referents. However, it should be emphasized that many of the indefinite expressions produced by the three- and five-year-olds in Expt. III occurred in nominative sentences, and cannot be taken as evidence of the ability to identify referents.

Maratsos (1974) has produced experimental evidence to suggest that, by four years of age, children understand the difference between a definite and an indefinite referring expression, and will use these expressions appropriately. However, Maratsos' experimental studies only cover a limited area of referential speech. Most of his evidence is based on question and answer sessions in which he told children a story, and then asked a *wh* question, to which the children were supposed to give an answer of the form 'the+noun' or 'a+noun'. The stories and questions were constructed to test three types of referring expressions: Task 1 – a referent is identified in the story (an X), and the child should therefore answer with an anaphoric definite reference (the X); Task 2 – a group of items are mentioned in the story (Xs), and the child's answer is required to identify one of them (an X); Task 3 – an indefinite, non-identifying expression is used in the story (an X), and the child should repeat this indefinite use of *a* in his answer (an X). Despite the ingenuity of this study, it is open to a number of criticisms. Task 1, testing children's ability to use anaphoric references, does not distinguish between children who refer because the referent has already been identified, and children who always refer, regardless of whether their referent has been identified. Task 2, testing children's use of identifying expressions, fails to distinguish between children who are using identifying expressions, and those who are merely naming. More generally, it may be argued that a speaker's ability to use the articles appropriately can be revealed most clearly when he is allowed to provide the verbal context for his referring expressions, rather than being constrained to respond to a verbal context imposed on him by an experimenter, as in Maratsos' studies. It seems quite possible that, by asking his subjects for a name in response to a question, Maratsos may have biased their responses in favour of the nominative indefinite article. Although it is not being suggested that the verbal context is the only contextual influence on referring expressions, nor that descriptive language represents all linguistic behaviour, the descriptive tasks employed in the present study would seem to be more generally representative of the referential use of language.

In conclusion, the experimental results reported above revealed some surprising contextual effects on the use of articles. In Expt. I, the task-order effect suggested that a subject's perception of the task and the function of his utterance may influence his decision to use a definite or an indefinite referring expression. Another odd result was the adults' failure to determine some of their referents in Expt. II, particularly in the Isolated condition. Perhaps this result may be explained as the poor enunciation of a relatively unimportant function word at the start of a sentence, but why then was there a difference between the conditions? And why did this effect not appear to the same extent in children's descriptions? It can only be suggested that some as yet unknown contextual factor, such as whether the listener is patently looking at the referent, or attending to the speaker, may be relevant here. It is of interest that children were less affected by these factors than adults. Clearly, referring expressions are influenced by context, both verbal and non-verbal. With this in mind, this exploration of children's use of the articles in referring expressions remains incomplete. In each of the above experiments, similar contexts were used; the referents were always in front of the speaker, and he was in face-to-face contact with his audience. The physical presence of both referents and audience may have encouraged the use of definite references, particularly by young children. It may be that young children would use identifying expressions more readily in contexts where either the referents are absent (e.g. recounting of a prior event) or the audience is absent (e.g. a telephone conversation). It also seems important to discover when children realize the dual function of the indefinite article, and whether the 'indefinite' sense of the article is mastered first.

This paper is based on a thesis submitted by the author in accordance with the regulations governing the award of the degree of Doctor of Philosophy in Psychology at University College London. The author acknowledges with thanks the invaluable advice provided by Dr P. N. Johnson-Laird during the course of this work.

REFERENCES

BLOOM, L. (1970). *Language Development: Form and Function in Emerging Grammars.* Cambridge, Mass.: M.I.T. Press.

BROWN, R. (1973). *A First Language: the Early Stages.* Harvard University Press.

CHRISTOPHERSEN, P. (1939). *The Articles: a Study of Their Theory and Use in English.* London: Oxford University Press.

GRUBER, J. S. (1967). Topicalization in child language. *Found. Lang.* **3**, 37–65.

MARATSOS, M. P. (1974). Preschool children's use of definite and indefinite articles. *Child Dev.* **45**, 446–455.

PERLMUTTER, D. M. (1970). On the article in English. In M. Bierwisch & K. E. Heidolph (eds.), *Progress in Linguistics.* The Hague: Mouton.

STRANG, B. M. H. (1962). *Modern English Structure.* London: Arnold.

VENDLER, Z. (1967). Singular terms. In Z. Vendler (ed.), *Linguistics in Philosophy.* Ithaca, N.Y.: Cornell University Press.

WARDEN, D. A. (1974). An experimental investigation into the child's developing use of definite and indefinite referential speech. (Unpublished Ph.D. thesis, University of London.)

(*Manuscript received* 4 *January* 1975; *revised manuscript received* 27 *March* 1975)

Br. J. Psychol. (1976), 67, 1, *pp.* 113–132
Printed in Great Britain

BOOK REVIEWS

Hemisphere function in the human brain. Edited by S. J. DIMOND and J. G. BEAUMONT. London: Elek Science. 1974. Pp. vii+398. £7.50.

This very useful collection of essays, many of them by authors well known in the field, offers an unsystematic but comprehensive survey of the anatomical and functional asymmetry of the human brain. The book as a whole allows the reader to reach a fair judgement of the current state of this heavily populated area of research.

If we compare this collection with the symposium on interhemispheric relations edited by Mountcastle in 1962, the most striking advance has perhaps been in anatomical discoveries. Geschwind presents an important review of the marked asymmetries that he and others have established in secondary auditory cortex, notably in the size of the temporal plane. He believes that these asymmetries provide the developmental origin of dominance. One difficulty, however, is that the anatomical asymmetries are found in only 60 per cent of adult brains, whereas the rate of right-sided dominance in dextrals, if we estimate it from the incidence of aphasia after unilateral lesions, is perhaps 3 per cent. It is interesting that the number of subjects who show asymmetry in a dichotic listening task is nearer to Geschwind's figure for anatomical asymmetry. The discrepancy might be explained if receptive functions were less lateralised than expressive functions: lateralisation of the analysis of speech may not be essential but it may be important that there is no competition for control of the articulatory organs.

Since 1962 there has been a large increase in the catalogue of functions attributed to the right hemisphere of the typical dextral. The list, reviewed in detail by Kimura and Durnford, by Dimond and Beaumont, by Levy and by Buffery, includes subitizing, stereopsis, the recognition of spatial position, the discrimination of complex forms in general and of faces in particular, and, provided the report is non-verbal, calculation.

Much of the more recent evidence for these and other asymmetries of function has come from visual experiments in which material is briefly presented in one or other hemifield. In this book the results of such experiments are almost always subject to a connectionist explanation: if presentation to one hemifield leads to more errors or to longer reaction times, then the conclusion is either that the material is processed less efficiently and more slowly by the recipient hemisphere or that the necessary operations can be performed only by the other hemisphere, whither the information must be transferred at the cost of degradation or delay. Of the authors in this book, only Levy considers the distinct hypothesis that has been advanced by Kinsbourne. And yet Kinsbourne's hypothesis explains many worrying aspects of the studies on perceptual asymmetry. In particular, it accounts for the lability of the effects, a lability that has not often been discussed in published papers but which is here treated very fairly by Kimura.

Like others, Kinsbourne supposes that the hemispheres are specialised for different functions, but his theory introduces an additional assumption. When the experimental task is such as to engage predominantly one of the two hemispheres, attention tends to be directed to the contralateral part of space. This bias of attention has its phylogenetic origin in the primitive turning reflexes and its consequence is that stimuli in the contralateral field are analysed sooner and more efficiently. Kimura and Durnford (p. 44) misunderstand Kinsbourne's theory and dismiss it in a brief final paragraph: '...a term such as "attention" is so vague as to be of little use in understanding the details of how the left and right hemispheres actually operate'. And yet Kimura and Durnford themselves discuss the dynamic nature of perceptual asymmetries and describe (p. 39) an unpublished experiment by Kimura that revealed a right-field advantage for the analysis of geometric forms provided that the form recognition was preceded by a task in which letters were to be identified. They write '...one hemisphere and/or a resultant scanning tendency can be "primed", so that stimulus material which ordinarily would not be asymmetrically processed, may be so processed as a consequence of the context'. It is not at all clear that their concept of 'priming' is less vague than the concept of 'attention' that they are to abuse a few pages later.

An attentional hypothesis will also account nicely for another finding by Kimura: subitizing of letters is better for the right half field even though the same task gave a left-field superiority when the elements were dots or geometric forms.

It would not be difficult regularly to introduce controls that distinguished between attentional and connectionist explanations of perceptual asymmetries. For there is usually a way of separating structural channels and perceptual space: it would be possible to present visual stimuli to the same retinal point when the eyes were deflected to one or other part of space; the apparent mid-point of the visual field could be displaced by prism adaptation; tactile stimuli could be presented to the crossed limbs; and the same auditory stimulus could be delivered to the two ears but with a phase difference. When the last of these was in fact done by Bertelson and Morais, the Kimura asymmetry effect survived. (Those who wished to retain a connectionist account of this result might, however, point to the observation of Evans that single units in the primary auditory cortex of the cat tended to be specific to positions in contralateral space rather than to the contralateral ear.)

It is also worth questioning the very legitimacy of the question 'What does each hemisphere do?'. Amongst psychologists rather than amongst neurologists there is a widespread view that the two hemispheres are relatively independent structures connected by a slender cable that has a low channel capacity; and that when the corpus callosum is cut we can examine each hemisphere performing its wonted function. This view is often detectable in the present book and becomes explicit in the chapter by Dimond and Beaumont, who regard the two hemispheres as two single channels: 'The analogy is not to a single computer, but to two computers sitting by side....' Yet the number of fibres in the corpus callosum, 180 million, is not of a very different order to the number of fibres with long axons within each hemisphere. It merely happens that this particular class of association fibre lends itself very readily to the surgeon's knife.

Pressing their concept of the two hemispheres as two single channels, Dimond and Beaumont argue that material is processed most efficiently if it is distributed between hemispheres. However, many of their experiments depend upon a strange apparatus that allows stimuli to be presented to either hemisphere by either eye, but which confounds hemifield with eccentricity: when the stimuli are presented to different hemispheres both stimuli are relatively close to the fixation point on 50 per cent of trials, but when they are both presented to the same hemisphere one of the stimuli is always many degrees into the peripheral field. There is no reason why both stimuli's being in the far periphery on half the trials should be equivalent to half the stimuli's being in the far periphery on all the trials. I earnestly hope that Dr Dimond and Dr Beaumont will abandon this unhappy instrument in the future. An experiment by Davis and Schmit, which Dimond and Beaumont quote in support of their argument, is open to an objection of the same class. Davis and Schmit presented two symbols either both to one hemisphere or one to each hemisphere and found that their subjects were faster in judging the identity of the signals if one signal was delivered to each hemisphere. But in the latter case the targets were separated by 20 degrees of visual angle whereas in the ipsilateral case they were separated by only 2·6 degrees. In all experiments in this field one must ensure that one is dealing with an effect of hemispheres and not with an effect of mere separation or position in visual space. For lateral interactions have been shown to be very extensive in peripheral vision: Grindley and Bouma have independently evolved the working rule that if two peripheral stimuli are not to interfere with each other then they must be separated by a distance equal to half their distance from the fovea.

However, elsewhere in the present volume, Trevarthen shows a wholesome scepticism in discussing Gazzaniga's early claim that commissurectomy increases a patient's capacity for parallel processing. Trevarthen indeed concludes that such patients are less able than normals to organize responses independently and simultaneously. He also presents a series of ingenious experiments that lead him to believe that the ambient perception of space is not divided by commissurectomy: thus two moving stimuli, one in each hemifield, can combine to produce a single complex percept in the split-brain subject.

Zangwill also reviews the results of commissurectomy and asks whether and in what sense the two hemispheres enjoy separate consciousness. He emphasises that these patients may show considerable deficits of attention and memory over and above the syndrome of hemispheric disconnexion that is usually described. As Zangwill suggests, it may be that these additional deficits are to be explained by the fact that the patients are epileptics, who are brain damaged in the first place (and whose brains, it should be added, are probably organized differently to those of normals); but it is also worth forgetting the anatomy, considering the brain as a unit and recalling again that perhaps two hundred million neurons have been destroyed.

The remaining chapters of the book complete the survey of the field. Levy discusses shrewdly and intricately the development of dominance and handedness. Pathological evidence for asymmetry of function is reviewed by Levy, Goldstein, Newcombe and Buffery. Simernitskaya and Hécaen and Marcie give detailed accounts of the types of disorder of writing that follow lateralized lesions. Gazzaniga describes some interesting attempts to teach aphasics a non-verbal communication system such as that taught to chimpanzees by Premack. He also suggests that 'cerebral dominance' should be taken to refer not to the lateralization of linguistic functions but rather to a single 'decision processing system that is responsible for bringing order for our various mental activities to the final cognitive path'. After the preference in the nineteen-sixties for the term 'cerebral asymmetry of function' we may perhaps expect a return to the useful metaphor of dominance.

J. D. MOLLON

Human Behaviour and Brain Function. Edited by H. J. WIDROE. Springfield, Illinois. 1975. Pp. ix + 116. $7.50 (U.K. price not given).

The blurb tells us that this Californian based offering is about 'advanced psychophysiological research in contemporary efforts to understand and treat a variety of behavioural disturbances' and that it is written by 'eight well known clinical researchers'. On this latter point your reviewer must shamefacedly admit to ignorance. None of the eight well known gentlemen were immediately familiar although a ninth contributor not included in the 'credits' (Ornstein) was.

The first chapter by Galin and Ornstein is a readable but not very penetrating overview of some of the Californian work on hemispheric specialization including reference to the split-brain work with humans. It might convince the casual reader that the rest of the book might also be worth looking at. Regrettably this is not the case and the subsequent contributions, with an occasional exception, deteriorate rapidly.

In chapter Two Bach-y-Rita presents a little homily to the effect that some individuals charged with crimes involving aggressiveness may be neurologically disturbed. The author of the next piece offers some speculations about drug effects and pleasure centres in the brain. These come from interviewing a large number of the junkies and other addicts with which California apparently abounds. This technique gets round the necessity to use the rather messy Olds and Milner business of sticking electrodes in rats' brains.

The undoubted *pièce de résistance* is the four pages by Voltolina and Widroe on the neurology of schizophrenia. This has all the brevity and punch of a letter to *Nature* announcing a major new scientific breakthrough. Electrical disturbances on the right side of the brain are associated with affective disorders and electrical disturbances on the left with schizophrenia. Schizo-affective disorders occur when the electrical disturbances on the right move across to the left hemisphere. One wonders why psychiatrists in general, and the rest of us concerned with mental illness, poor demented lot that we are, have not latched on to this simple truth years ago?

Although there are one or two chapters which make a much better showing and which were probably quite appropriate for the setting in which they were originally presented, the plain truth is that this symposium given at the Walnut Creek Psychiatric Hospital did not merit being recorded for posterity in this way. It is not 'thought provoking and pragmatic for clinicians and researchers' in quite the way that the blurb writer meant and it would have been better if it had remained quietly 'up the Walnut Creek'.

EDGAR MILLER

Psychology and Medicine. By S. J. RACHMAN and CLARE PHILIPS. London: Temple Smith. 1975. Pp. 223. £3.75.

Drs Rachman and Philips have written an extremely readable and important book which attempts to achieve two aims. The first of these is to make out the case that psychology is relevant to medicine and the second, more implicit aim, to indicate possible future directions in clinical psychology.

Some hard headed readers might be put off by the thought of yet another book about psychological aspects of medicine. Many of these works are either psychodynamic treatises, or a kind of Noddy psychology urging doctors to be nice to patients. The present work falls into neither of these categories. It is evidence based and covers the actual and potential contributions of psychology in areas such as pain relief, sleep disturbances, possible uses of biofeedback techniques in the treatment of physical illness, and headache, as well as in more traditional fields such as subnormality and mental and behavioural abnormality.

As it is an introductory work there has necessarily been a high degree of selection, but amongst the examples chosen are the toilet training programme of Foxx and Azrin; Janis's studies of preoperative anxiety and post-operative adjustment; Stacey's research on the effects of hospitalization on children; Stuart's self-control procedures for the obese; Budzynski's biofeedback management of headaches; and Shapiro's research on biofeedback control of blood pressure. Some of these examples are of course on the frontiers of knowledge and the authors are appropriately cautious when discussing them. But it should also be of interest and value to anyone with an interest in this field.

Doctors and psychologists should find much of value in this book, and because of its readable style and interesting content it will be a required addition to the reading lists of courses on psychology for medical students. It should do much to overcome the medical students' frequent complaint that psychology is irrelevant to medicine.

Clinical psychologists who are thinking about the way their profession should develop will find guidance here. Indeed this introductory work makes out its case so well, that we now need a larger work which gives in-depth treatment of the topics covered – a handbook of non-psychiatric clinical psychology!

PHILIP LEY

Anxiety Factors in Comprehensive Patient Care: Where General Medicine and Psychiatry Meet. Edited by W. LINFORD REES. Amsterdam: Excerpta Medica, 1973; and New York: American Elsevier, 1973. Pp. v+106. $10.40.

Readers will have marvelled often at the wealth behind the medical profession which facilitates arrangement of and attendance at symposia and at the arrogance which encourages publication not only of the papers but also of the trivia, repetitions and contradictions of 'Discussion' as if those who could not attend must not be deprived of any pearl of wisdom.

In this volume – Proceedings of a Symposium on 'Anxiety Factors in Comprehensive Patient Care' sponsored by Wyeth International and held in Amsterdam in March 1973 – the problems of such publication are exacerbated by lax editing (which gives us, for example, 'the cultural nervous system' p. 41) and one looks in vain for the pearls. Granted there is food for thought for general practitioners and material here for discussion between psychologists and medical colleagues. This is especially true of the contributions of van Beusekom (on the anxieties of general practitioners), Meltzer (on cardiovascular disorder and anxiety) and Bruyn (on pain). But the chapters are almost uniformly noteworthy for their superficiality, their lack of originality and their bland statement of opinion without evidence – particularly their uncritical acceptance of the value of psychotherapy.

However, it is in the current climate of enthusiasm among psychologists and, we observe, among physicians for mutual collaboration that this book is most remarkable and most disappointing. Nine international symposiasts on the topic of anxiety in general medicine include a general practitioner, a cardiologist, a neurologist, a neurophysiologist, two specialists in obstetrics and gynaecology and three psychiatrists. Thus every speaker was medically qualified. Not only was there no psychologist and not one speaker felt called upon to deplore or even comment on this gap in representation, but, with the notable exception of Bruyn, they all managed to discuss anxiety, often with special reference to its definition, assessment and treatment, without more than a passing reference to any psychologist or to psychology as such. Behaviour therapy is mentioned but never referred to its sources in psychological theory and clinical practice. Only Bruyn outlined J. A. Gray's theory of the structure of the emotions and the limbic system. Otherwise the extensive psychological theories of emotion and of anxiety in phobic states, in compulsions, in schizophrenia, in depression and in normal behaviours might not exist as far as these medical experts are concerned.

Now Linford Rees, the Chairman and editor, has had the advantage of working with some of the most eminent psychologists of today. His writing here shows that he *has* learned from psychologists. Why, then, we are driven to ask, is there this studied avoidance of psychology in an area so central to psychology and the movement for psychologist/physician cooperation? The question is particularly relevant since it is raised also from another angle by Rachman and Philips' (1975) plea for development of collaborative schemes. Perhaps, despite their efforts to the contrary, their proposals may appear trifling and naïve to medical men. Equally, the lack of philosophical and psychological sophistication of the Amsterdam symposiasts will jar on psychologists. But Rachman and Philips are at least calling for a reunion. Linford Rees and co. seem to be avoiding this possibility.

Several hypotheses could be put forward to account for this phenomenon. For example, the symposiasts could be as truly ignorant of psychology as their unscholarly publication suggests. As we have seen, this cannot be true of their Chairman and hence is an unlikely hypothesis. Secondly, medical specialists could be scared, as psychiatrists are said to be, of psychologist take-over. Given the magnitude of behaviour problems in the community and the paucity of professionals to deal with them, this has never been a convincing argument within the domain of psychology and psychiatry. It has still less cogency in relation to other medical specialities. However, there may yet be some substance in this hypothesis. Thirdly, the medical profession may be so inward-looking that it will minimize views and evidence from other sources than medicine. There is some support for this hypothesis even within this slim volume. Throughout, only two psychological journals, rather than psychiatric or medical, are listed as references. If then this hypothesis is confirmed, we may be forced to conclude that psychologists will not succeed in collaborating with physicians to any great extent until we have powerful advocates *within* the profession of medicine. Rachman and Philips refer to clinical psychologists who may be physicians manqués. Such may exist but they are well disguised or, inevitably, they are poor spokesmen for our profession. Equally, there are many able experimenters among the psychiatrically qualified who might properly be thought of as behavioural scientists, and, there have been those who have taken both medical and psychological qualifications but, to date, they speak still from the fold of medicine. The practice of medicine is a way of life but so, too, is the study of psychology. We await the individual, or better still the group, who will have the theoretical and financial courage both to obtain dual qualifications and then work as psychologist and for psychology, speaking to medical colleagues from within their profession and convincing them of the value to them and their patients of objective, scientific behaviour study.

In the publication of a book under this title we looked for the spearhead of this revolution within medicine. Psychiatry, as Rachman and Philips point out, has pioneered by introducing psychology to behaviour disorder. Psychiatry may yet point out the advantages of psychology for general medicine. But on this occasion a golden opportunity has been lost. We must yet make our own efforts and also look to the young doctors and, perhaps, the physician/psychologists to show the place of psychology in the study of anxiety and other psychological factors in comprehensive patient care.

ANNE BROADHURST

RACHMAN, S. J. & PHILIPS, C. (1975). *Psychology and Medicine*. London: Temple Smith.

The Psychophysiology of Mental Illness. By M. LADER. London: Routledge and Kegan Paul. 1975. Pp. x+270. £7.50.

This is the second volume in the series *Social and Psychological Aspects of Medical Practice* and was written with psychiatrists, clinical psychologists and general medical practitioners in mind.

After a brief introductory chapter in which the author worries, perhaps unnecessarily, about what exactly constitutes the proper area of study for psychophysiologists, there is a long (72 pages) chapter on psychophysiological techniques. The survey of techniques is thorough, extending from brief descriptions of such instruments as thermistors and photoelectric transducers and the like, to concise and extremely readable accounts of electroencephalography and endocrinological assay methods. The techniques chapter should serve equally well as a first introduction to the area, to be read through from start to finish, or as a source of information to be dipped into as the need arises. What, for example, *are* the adrenal steroids found in the urine, how are they estimated and why? Look it up in this chapter and you will be provided with a quick outline of the topic together with a few basic references should you wish to follow the topic up any further.

The next chapter, on some psychophysiological concepts, is not nearly so useful. The various concepts selected for coverage (stress, biofeedback, activation, and so on) are dealt with rather too sketchily, as though emphasising the author's major concern with techniques rather than their theoretical implications, which I am sure is far from being the case.

Three chapters then follow each of which is devoted to a specific diagnostic category – anxiety, depression, and schizophrenia. In each of these chapters Dr Lader does a difficult job well. By combining the styles of a literature review and a teaching text he manages to produce an account which is both authoritative and readable. Anyone reading these chapters, and the next one on miscellaneous conditions such as hysteria, depersonalization, psychopathy, and psychosomatic conditions, should not only obtain a thorough grasp of the present state of psychophysiological knowledge in these areas, but also have no difficulty in following future developments in the journal literature.

Were it not for the rather daunting price (maybe a paperback edition is planned?) this book could easily find its way into undergraduate teaching courses where it would serve admirably the dual functions of dispelling the mystery which surrounds psychophysiology for many students, and inspiring others to take a deeper interest in those areas currently classified as 'biological psychiatry'.

If future volumes in this series are as good as this one we may look forward to a well informed future generation of clinical psychologists and psychiatrists and possibly (who knows?) maybe even general practitioners.

F. N. JOHNSON

Abnormalities in Parents of Schizophrenics. By S. R. Hirsch and J. P. Leff. London: Oxford University Press. 1975. Pp. 200. £5.50.

This excellent book is the twenty-second in the Maudsley Monograph series and bears resemblances to many of its predecessors. The book is in two sections: an extended review covering the literature on the parents of schizophrenics and a report of the authors' own investigation of one particular set of hypotheses.

The literature review is long by any standards, occupying six chapters and well over half of the book. Hirsch and Leff examine dozens of studies of the parents of schizophrenics under the headings: questionnaire studies, small group interaction studies, abnormal thought processes, abnormalities of communication and language. Although the review is clear and always critical on methodological points it tends to develop into too much of a catalogue to be really absorbing reading. There is no doubt however that this review, perhaps supplemented by Jacob's (1975) extended review, will tell anyone all they want to know about the empirical research on the families of schizophrenic patients.

The major problem with many such studies is that they can actually tell us little about the nature of any link between family interaction and schizophrenia. It is sometimes assumed by researchers and textbook writers that the demonstration of the existence of abnormalities in parents of schizophrenic offspring implies a cause and effect relationship. Hirsch and Leff repeatedly point out that such a link may equally well be evidence of a common genetic endowment more fully expressed in the patient, or the response of the parents to the difficulties involved in raising a psychotic child, as evidence of a schizophrenia producing social environment. Hirsch and Leff are careful to avoid this mistake in the report of their own studies.

At the conclusion of their review the authors conclude that several hypotheses about the parents of schizophrenics have been supported by reasonably good evidence. These include the presence of schizoid mothers; conflict and disharmony between parents; the presence of physical illness among pre-schizophrenic children; schizophrenic's mothers are more concerned, protective and intrusive than the mothers of normals; and possibly that the parents of schizophrenics display abnormalities of communication.

It is with hypotheses of this last kind that Hirsch and Leff's own study is concerned. They take the findings of Singer and Wynne at NIMH as their starting point and attempted to replicate these findings in London. It is Singer and Wynne's belief that communication 'deviances' in parents are a contributory *cause* of schizophrenia in children. Hirsch and Leff are commendably more cautious and just look for the existence of a relationship and leave speculations about cause and effect alone. Based on parents' responses to a Rorschach Test, Singer and Wynne found that they could blind rate which subjects had schizophrenic offspring with remarkable accuracy. They looked for such things as 'gross indefiniteness', 'contradiction', 'nihilistic remarks', swearing, odd sentence construction, reiteration and *non sequitur* reasoning and found these to be far more prevalent among parents of schizophrenics than among parents of neurotics or normals. In their replication Hirsch and Leff found more such instances of deviant communication among parents of schizophrenics than neurotics but to nothing like the same extent as their American counterparts. In fact they found a 95 per cent overlap in deviancy scores of schizophrenics' and neurotics' parents ($n = 80$) whereas Wynne and Singer found no overlap at all.

After an extensive and thorough analysis of all their data Hirsch and Leff are left with two questions, why did they get the result they did and why is it so discrepant with the American research. After examining a number of alternative explanations they are left with the hypothesis that the *fathers* of schizophrenics are more verbose and marginally more deviant in communication than fathers of neurotics. They do not see their results as supporting Wynne and Singer's hypothesis that communication problems are universal and unique among schizophrenics' parents however. The reasons given for the differences in the outcome of the two studies relate to differential diagnostic and referral practices on both sides of the Atlantic.

The book is thoroughly researched and well presented and the study contained in it is thoughtfully designed and meticulously carried out with unusual and commendable attention to testing and scoring procedures. Most readers will be far more inclined to trust the results of this study than the earlier one. However whether the study warrants book length treatment is debatable and at today's prices perhaps a book is too extravagant. There appears no real reason why the review and the research report could not have appeared separately as journal articles of more moderate length. However the material does appear in book form and is certainly worth reading.

RAYMOND COCHRANE

JACOB, T. (1975). Family interaction in disturbed and normal families: a methodological and substantive review. *Psychol. Bull.* **82**, 33–65.

Psychiatric Diagnosis: a Review of Research. By G. FRANK. Oxford: Pergamon Press. 1975. Pp. 140. £5.40.

The author of this expensive little book set out to determine, by reviewing the literature, what evidence there was for the usefulness and validity of psychiatric diagnoses. The greater part of his text is an exhaustive description of the innumerable factor analytic studies of patients' symptom ratings and test scores carried out, mainly by American psychologists, in the 1950s and 1960s. At the end of it all – there are over eight hundred references occupying more than a third of the book – he is driven to conclude that 'psychiatric diagnosing serves no practical purpose' and should be abandoned forthwith. 'What we need in the study of psychopathology', he tells us, 'is a revolution', though it is not clear who is to lead us, or in what direction.

Fortunately or unfortunately, depending on one's viewpoint, these drastic conclusions do not follow from the evidence presented. Factor analysis of symptom ratings has proved unrewarding not because Kraepelin's diagnostic categories were ill chosen but because factor analysis is an inappropriate statistical technique either for validating an existing typology or replacing it with a new one. It is a means of defining dimensions of symptomatology, not of identifying clusters of like individuals. There is no reason why the number of factors obtained should correspond to the number of sub-populations present, or why their profiles should reflect the identity of such populations. Indeed, as the algebra of factor analysis is based on an assumption of multivariate normal distribution, the crucial issue is begged from the beginning.

The sad fact is that the greater part of the uninspiring literature Dr Frank has so painstakingly reviewed is irrelevant to the question he set out to answer. The validity of psychiatric diagnoses depends primarily on their predictive value; whether, that is, there are consistent and significant differences in response to different methods of treatment and long term outcome between one category and another. The answer to these questions depends on clinical trials and longitudinal studies of representative cohorts, not on factor analysis. Other multivariate techniques such as cluster analysis may be invaluable for identifying new and unsuspected groupings of like individuals, and discriminant functions may help to validate existing diagnostic distinctions, but neither of these techniques is mentioned by Dr Frank. Nor does his brief chapter on outcome studies pay any attention to the vast literature on the differential response of different syndromes to phenothiazines, lithium, tricyclic antidepressants and monoamine oxidase inhibitors. We are simply told that diagnosis is no use in the prediction of drug response, and given a single reference eight years old.

Psychiatric diagnoses do indeed have many shortcomings. They are often unreliable, they are sometimes misused so that they become a barrier rather than an aid to effective treatment and understanding, and their validity has been taken for granted far too long. But we have to classify our subject matter before we can communicate with one another, or even with ourselves. To discard a classification out of impatience with its imperfections before something demonstrably better is available would simply be to invite chaos. What we need is not a revolution, but a little more thought.

R. E. KENDELL

Modern Perspectives in the Psychiatry of Old Age. Edited by JOHN G. HOWELLS. Edinburgh: Churchill Livingstone. 1975. Pp. 630. £10.00.

'The 81-year-old confused and incontinent woman, whose delusion is that she is pregnant, clings to what is to her a reality despite all objectively true findings to the contrary. For truth is objective, while reality is subjective, varying with the individual and becoming clear to anyone willing and able to interpret its meaning. What the old woman was saying was that she harbored her own regressed self within herself and was about to deliver it. If senility is the "second childhood", then she was pregnant with it and at the point of delivery. Viewed from this perspective, her delusion makes sense, her method of expressing it, poetry.'

Most of the other contributors would, to judge from their papers in this *Modern Perspectives*, Volume no. 6, have interpreted this psychopathological phenomenon somewhat differently. There are 28 chapters and 35 authors who are mainly of Anglo-American background, but there are also two Russian, one West German, one Japanese and one Finnish worker. The quantity and quality of these contributions demonstrate the astonishing growth of a subject whose potentials came to be recognized only some twenty-five years ago. All aspects of the psychiatry of old age are fully covered. In fact, in some areas like psycho-pathology, functional illnesses, suicide and various forms of care, there is considerable repetition and overlap, especially as some contributors were not editorially discouraged from going beyond their allotted task. For example the chapter on 'Integrated Services' is introduced by a long account of clinical tests of cognitive dysfunctioning favoured by the author.

The book has thus become rather bulky and for this reason alone could not be recommended to post-graduate students. However, as the editor himself writes, textbook coverage is not attempted in the Perspectives series. Most chapters have up to 100 or more references which are gratifyingly up-to-date including publications during 1973. Psychiatrists, psychologists, nurses and social workers wishing to inform themselves on some recent developments are likely to find the information in this volume. Chapters of specific interests to psychologists are those dealing with the pathology of dementia (103 references), the psychology of ageing (191 references), mental retardation (a new area of inquiry in old age, 19 references), and psychometric techniques (116 references).

F. POST

Psychogeriatrics: An Introduction to the Psychiatry of Old Age. By BRICE PITT. Edinburgh: Churchill Livingstone. 1974. Pp. vii+183. £3.00.

This little book presents a concise account of the problems associated with old age and, in particular, the psychiatric illnesses of this group. It does so with skill and refreshing simplicity of style.

The author, Dr Brice Pitt, although dealing with the subject in a scholarly way, giving valuable series of references at the end of each chapter, brings a lightness of touch to his theme which illuminates the whole book. He lays particular emphasis on the importance of recognizing the psychological and environmental pressures upon the elderly and suggests social as well as medical strategies to deal with them.

The first three chapters deal with ageing and its general problems and professional attitudes to them. These are followed by a series of more clinical chapters on dementia, depression, mania, paranoid states, neurosis and personality and behaviour disorders, seen in the elderly. Emphasis is laid on the importance of accurate diagnosis and appropriate treatment in contrast to the therapeutic 'nihilism' sometimes seen in the care of the old.

The chapters on 'A Psychogeriatric Service' and 'Principles of Treatment' lay stress on the value of an assessment team made up of a consultant psychiatrist, nurses, social workers and occupational therapists, but there is one curious omission, that of the psychologist. On page 40 on the chapter on 'Dementia', psychological testing is referred to as 'of little help in diagnosis, for no tests wholly distinguish brain damaged patients from those who, for other reasons, are too perturbed or preoccupied to co-operate'. Further on, the author says that 'the psychologist can map out the patient's disabilities more precisely than can the psychiatrist at an ordinary clinical interview and, by successive testing, can show in what ways there is any improvement and deterioration'. The psychologist in fact has an important function in the assessment of all elderly patients. The failure to emphasise the value of the psychologist in this field of work is perhaps the only criticism which could be levelled against this otherwise excellent book.

It is to be recommended particularly to those who have thought the study of old age to be a dreary task. Their conversion should be rapid and, in the majority of cases, permanent!

ISABEL MOYES

Memory, Related Functions and Age. By JACK BOTWINICK and MARTHA STORANDT. Springfield, Illinois: Thomas. 1974. Pp. vii+198. $12.75.

The study reported in this book aims to sample a variety of aspects of memory, and closely related behaviour, over the age range 21–80 years. Although this is a very laudable aim, the sample size of 120 is far too small when such broad questions are being investigated, especially as the men in the sample only completed half of the tests of memory related behaviour. However, with some reservations, the work is none the less worthy of attention by those in this field.

A battery of tests were given to the 120 volunteers drawn from the St Louis area. The memory related functions included tests of brain function/perception (e.g. Hooper's Visual Organization Test and Trailmaking), psychomotor speed (e.g. crossing off and copying digits), the WAIS vocabulary scale and various questionnaires and rating scales covering health, satisfaction and habits of drinking and smoking. The memory tests ranged from digit spans to paired associate learning, through visual recognition and recall to memory for historic events. Unfortunately, in both groups of tests measures were included whose reliability, and validity, must be seriously doubted.

The book is neatly split into two sections, memory related functions (Part I); and the various measures of memory themselves (Part II). Each of these sections is terminated by a factor analysis of the results presented in the foregoing chapters, and at the end of the second section there is a factor analysis of the combined results of Part I and Part II.

The presentation of the work reported is very empirical; very seldom is the reader lifted above the bare results and shown the broader theoretical issues. There is a chapter by chapter presentation of individual results and at times this becomes a little laborious. When the results of factor analysis are presented, there are attempts to link the findings with a few other studies and with theoretical positions, but these attempts are weak.

In fact the findings are largely in accord with other published work. The commonly reported age declines in performance are shown on tests of psychomotor speed, visual organization and block design and, of course, most aspects of memory. Three psychologically meaningful factors are extracted from the memory data. One factor seems to represent general short term memory, with loadings on a wide range of tests; another is memory for spans and patterns, which the authors interpret to be a more immediate form of memory than the first factor; and thirdly, a factor of long term memory, loading on little else except the tests of LTM. The first two factors show age related declines in performance, but the LTM factor does not.

When the combined memory, and memory related behaviour results are analysed, the same factors appear, along with a fourth meaningful factor, a factor relating to health and personal outlook. From the memory related data the test of visual organization loaded heavily on general short term memory, whilst the tests of psychomotor speed loaded heavily on the factor for spans and patterns. Botwinick and Storandt suggest that this latter factor could in fact represent ability to deal with rapid sequential processing of information, whilst the general STM factor is more dependent upon registration, storage and retrieval of information. It would have been valuable for these ideas to have been more fully developed and placed in a broader theoretical context.

This book is valuable to those working on memory, particularly memory in the aged, but it will do little to excite the interests of the more casual reader. ANDREW J. GIBSON

Studies in Long Term Memory. Edited by A. KENNEDY and A. WILKES. London: Wiley, 1975. Pp. xviii + 358. £10.00.

Historically, British psychology played an important role in the development of research on long term memory, through the pioneering work of Sir Frederic Bartlett. In more recent times, however, psychologists in this country have been more attracted to the study of short term retention and have left long term memory as the playground for the practitioners of paired-associate learning. Now, with the Ebbinghaus empire in full retreat and Bartlett's ideas on the active, constructive nature of memory once more to the fore, it was entirely appropriate that the SSRC should sponsor a major conference on the content and structure of long term memory. This was held in Dundee in 1973 and brought together psycholinguists, clinical and experimental psychologists to report upon and discuss the latest findings in the field. Revised versions of thirteen of the original papers, together with four new articles, have now been published in book form, edited by the convenors of the original conference.

The editors have attempted to impose some structure upon the rather diverse contributions by grouping them under five headings and providing an introductory integrative chapter. While their efforts are commendable, the material still suffers from that lack of focus which so often characterizes collections of Conference papers. Individually, however, the various contributors have much to offer.

The first section on 'encoding' opens with the almost obligatory article by Broadbent, this time attempting to demonstrate, from a rather shaky empirical base, that the size of the basic memory unit is not five or seven, but three items. Wilkes also gets embroiled in the numbers racket in his interesting account of how cognitive structure can be inferred from subjects' pauses while reading and recalling word lists. In the same heterodox section, Schulman considers some of the wider implications of the 'depths of processing' approach to memory while Paivio contributes a useful update to the material reviewed in his *Imagery and Verbal Processes*.

To my mind, the highlight of the second section on 'organization' is the elegant critique of theories of semantic structure in general and those of Collins and Quillian in particular, offered by Johnson-Laird. Kiss demonstrates how, with the aid of a computer and a large number of patient subjects, it is possible to 'grow' an associative network which simulates some of the properties of human thought. A useful note of caution is struck in the same section by Morton who emphasizes the importance of separating the data base, from the way in which the individual subject chooses to operate upon it for the purposes of memory retrieval.

The section on 'integration' is, again, rather unintegrated. Kennedy, in a difficult but rewarding article, examines the age-old problem of how when reading prose, we turn a series of discrete words into a continuous flow of meaning. Levelt and Kempen, on the other hand, provide a stout defence of the traditional approach to sentence comprehension in terms of syntactic structures, utilizing much unpublished European research.

Under 'retrieval' Seymour presents a closely argued case, based on his own experimental work, for the view that visual and verbal processes are subordinated to some higher-order abstract coding system. Those who attended the original Conference will be sad to see that the paper by Meyer included in this section is not his original sparkling contribution, which has since been published elsewhere. Nevertheless, the new paper, which presents a model for processing affirmative and negative sentences, is not without interest. The book concludes with a section of the pathology of memory. Marshall's provocative paper on hemispheric specialization in the processing of parts of speech finds a home here, as does Baddeley's lucid exposition of theories of amnesia.

As this incomplete summary suggests, the book contains a great deal of interesting material, but it is hardly to be recommended as a comprehensive survey of research into long term memory, *c.* 1973. The editors have quite deliberately emphasized the psycholinguistic approach to the structure of long term memory which stresses the importance of the retrieval of semantic information for effective comprehension. The more orthodox function of long term memory, as a repository of events or experiences stored over time, receives much less attention. Thus, the results from the free recall paradigm, probably the most extensively used tool in contemporary research into long term retention, get scant coverage. Another notable omission is any treatment of the developmental aspects of memory function, other than a paper on the rather peripheral topic of skill integration. Readers who wish to know more of the large literature on what Tulving terms 'episodic' memory as opposed to the 'semantic' memory studies highlighted in the present volume, may readily consult *Organisation and Memory*, another set of conference proceedings, edited by Tulving and Donaldson.

The present volume will readily find a place on the specialist's bookshelf but the rather selective coverage of topics, combined with the high level of background knowledge required to appreciate some of the more technical papers, are unlikely to endear it to a more general audience. Both groups will be appalled at the price.

GRAHAM DAVIES

The Cybernetics of Human Learning and Performance. By G. PASK. London: Hutchinson. Pp. 347. £7.75.

Psychologists who have noticed of late the reduced currency of the term cybernetics, except in the Dr Who television series, will undoubtedly show interest in the latest contribution on the topic by Professor Pask, one of our more seminal thinkers. The title seems confidently to offer a reassertion of the central role of cybernetics in two important areas of psychological research. However, all is not as it may appear. This is a book for the cybernetician, the logician, the mathematician, the linguistic analyst and the electronics engineer but not, I fear, for the psychologist. It is written in a manner which few psychologists will find easy to follow and it fails to answer many of the questions which they will want answered concerning the possible contribution of the cybernetician's art to the psychologist's problems.

In my opinion the book fails mainly at the simple level of communication. This failure does not result from the inherent difficulty of the material dealt with since many of the topics such as information transmission and finite state automata are explained lucidly enough in other texts. Instead the problem appears to lie in the author's prose style and his manner of organizing the material to be presented. There are many confusing failures of syntax and punctuation, continuous misuse of brackets within a sentence to express ideas which could easily follow in a supplementary sentence, and persistent refusal to summarise ideas or give due warning of what material is central to the author's theme and what is merely qualifying detail. Words are regularly given meanings which would give even Humpty Dumpty pause: 'Hence, the evidence that a *concept* of something exists and is reconstructible, permanent or memorable is an explanation *of* an explanation; which (if properly attested) is called an *understanding*' (author's italics, p. 280). This I read to mean that, if I can say how I solved the problem, then I can be said to have understood it. I may be wrong in my interpretation – and probably am – but, if so, it is as much the author's fault as mine. It may seem unfair to attach so much importance to the stylistic deficiencies of the book but, if you attempt to read it, I am sure that you will rapidly come to sympathize.

Setting aside the obvious irony of the failure of a book dealing with aspects of communication theory to communicate, I found little in the book that was both novel and interesting. The preliminary discussion of the nature of information and finite state automata, despite some brilliant footnotes and intriguing asides, was heavy going and full of hair-splitting detail whose relevance, if any, was never explained. The interesting possibility of 'fuzzy abstract machines' was raised but their potential never fully assessed. Controversial issues such as the viability of the information measure in psychology or the defensibility of the limited-capacity, single-channel conception of man's cognitive capacities were never even raised, let alone satisfactorily dealt with.

Much of the subsequent discussion of training techniques reflected considerable sensitive insight into the learning process in practical situations. However I could detect little that was novel. Moreover, most of the discussion was in the context of very artificial, highly abstract learning tasks involving pushing combinations of buttons in response to combinations of lights which did not encourage generalization beyond such tasks which might, in any case, better be performed by simple finite state automata. The general drift of chapters 7 and 8 was that tasks are learnt most efficiently if the training exercises were maintained at a level neither too easy nor too difficult. I find it difficult to disagree.

Attempts in chapter 10 to describe human action control in the language of serial processing computers ran into difficulties which were only partially overcome by permitting parallel processing devices. The approach was, however, almost completely abandoned in chapter 11 when the book suddenly came alive with the study of a realistic problem in a classroom environment. This experiment was the only recent investigation I was able to detect in the book and, despite some reservations concerning his bold dichotomy of learning strategies into holist (*sic*) and serialist, well worth getting to grips with.

It would be wrong to reject out of hand a three hundred page contribution on the topic by an internationally acknowledged expert, especially since there are many flashes of wit and insight – instances where the hand of the master is clearly visible. Nevertheless as a psychologist interested in cybernetics, human learning and performance I found the book exceedingly heavy going and I resented the time and effort spent wading through the confusing prose to uncover these rare, if precious, nuggets.

RAY MEDDIS

Attention and Performance V. Edited by P. M. A. RABBITT and S. DORNIC. London: Academic Press. 1975. Pp. xviii + 732. £16.

This book maintains the tradition of earlier volumes in this series by presenting the viewpoints of several contributors on a number of disparate topics. This time there are forty-four contributions, which are rather loosely organized under the five headings of Attention and Set, Immediate Memory, Information Transmission, Long Term Memory and Conceptual Organization, and Perception and Comparison of Complex Stimuli. As may be surmised from these headings, a reader deriving full benefit from this book would need to be a divergent-thinking polymath, a status which your humble reviewer does not claim for himself.

In general terms, the book is a frustrating amalgam of the good and the bad. On the positive side, it contains a number of excellent theoretical and experimental contributions which definitely further our understanding and knowledge of various topics. In addition, it is unusual in providing a genuine feel of the research process, with several contributors presenting data which they believe to be important, although as yet they are unable to integrate these data into a coherent conceptual and theoretical framework. This public willingness to avoid premature closure makes a refreshing change from the somewhat arid theoretical certitude of textbooks in psychology. On the negative side, reading this book makes considerable demands on the reader, as the Preface disarmingly admits. However, for a volume of this length, diversity, and price, it is little short of scandalous that it is totally bereft of indexes. While the brilliant and knowledgeable may effortlessly detect interconnexions among the various contributions, the rest of us would have been considerably aided by the author and subject indexes that were, after all, provided with *Attention and Performance IV*. Furthermore, *Attention and Performance IV* was provided with extremely useful tutorial reviews, which provided a general orientation for the non-expert, but these, too, are missing from the current volume. Thus the non-expert, confronted by a heterogeneous collection of chapters containing data whose significance is unclear to the expert contributors, and denied crutches either of the index or review variety, is likely to come away from the volume with his mind confused and befuddled.

Interesting developments in work on attention are reported by Shiffrin and by Moray. Shiffrin discusses several experiments in which a number of input channels had to be monitored for a specified target. Under the simultaneous condition, the subject was required to monitor all channels simultaneously, whereas, under the successive condition, the subject could devote all his attention to each of the channels in turn. Perhaps surprisingly, Shiffrin found equivalent performance under both conditions, leading him to suggest, along lines initially proposed by Deutsch and Deutsch (1963), that selective attention affects processing in short-term store, but does not affect initial, perceptual processing. Since, at the most, a single signal occurred on each trial, one might feel inclined to argue that these results are probably only obtainable where the information-processing load is low. However, Shiffrin rejects this argument for reasons which are unclear. Moray discusses data and theory which, in spite of the different terminology, seem in sympathy with the Shiffrin approach.

In the section on immediate memory, there is an interesting chapter by Baddeley, Grant, Wight, and Thomson. They investigated the effects of imagery on learning by using interference tasks. While the results were not definitive, they do indicate the possibility that there is a fundamental distinction between concreteness, which is seen as an important characteristic of a word's representation in semantic memory, and imagery, which is seen as an activity or process. The prime evidence for this dichotomy is that retention of concrete and abstract words was not differentially affected by concurrent visual activity, whereas the use of imagery led to disruption by concurrent visual pursuit tracking. Crowder makes some perceptive comments about the inadequacies of several of the studies attempting to demonstrate the existence of echoic memory, and Shallice defends the traditional position that primary memory is concerned with phonemic but not semantic information, in spite of the evidence of Shulman (1970).

The section on long-term memory contains a number of useful contributions. Morton and Byrne report some preliminary data on retrieval viewed as problem solving. As a few of us have suspected for some time, the simplistic theories of retrieval currently popular have little to tell us of the processes involved in, for example, generating shopping lists or suggesting the menu for a meal. It is high time for the problem of retrieval to be researched in paradigms other than that of free recall. Mandler continues to impress upon us the fact that several studies indicate that retrieval from long-term memory is apparently limited to about five items at a time. While one might assume that this limitation stems from the limited capacity of short-term store during retrieval, Mandler argues somewhat obscurely that the limitation is one of storage. While his theoretical ideas are debatable, Mandler presents some fascinating data. For example, he asked subjects to name common objects to be found in an office, and observed an embarrassing pause after four to six items which he refers to as 'experimental aphasia'. Smith and Jones, in a thought-provoking chapter, discuss the possibility that semantic memory is more similar to a dictionary than conventional wisdom allows, and that alphabetic, or phonemic, structure is of importance in retrieval of information from semantic memory. The suggestion is a plausible one, and while the experimental evidence they present is somewhat sketchy, their emphasis on phonemic aspects of semantic memory is a useful counterweight to the prevalent focus on semantic aspects.

In the final section of the book, Baron continues his studies into various aspects of word recognition. At least under certain conditions, he has found that orthographic, rather than phonemic or semantic factors, are of prime importance in determining performance in word-recognition tasks. Broadbent and Broadbent found that inability to recall an item did not necessarily impair its recognizability. This finding is at variance with conventional wisdom as well as with most theories of recall and recognition. Broadbent and Broadbent make the plausible suggestion, based on their subjects' reports, that recognition involves the retrieval of extra-list information, whereas recall necessitates the retrieval of intra-list information. Corballis provides a good review of work within the Sternberg tradition, and correctly suggests that some synthesis of exhaustive-scan and direct-access hypotheses is in order.

I have tried, from my own undoubtedly biased viewpoint, to highlight some of the theories and findings which I found to be of especial interest or significance. As will be appreciated, the book is like an appetizing stew; unfortunately, the cooks have left it to the diner to stir the ingredients together.

MICHAEL W. EYSENCK

DEUTSCH, J. & DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.* **70**, 80–90.

SHULMAN, H. G. (1970). Encoding and retention of semantic and phonemic information in short-term memory. *J. verb. Learn. verb. Behav.* **9**, 499–508.

Cognitive Organisation and Change: An Information Processing Approach. By ROBERT S. WYER, JR. Chichester: Wiley. 1974. Pp. x+502. £10.50.

We learn in the preface that the book derives from a graduate course taught by the author in cognitive organization and change. It attempts primarily to provide a general conceptual framework for analysing research and theory on attitude and opinion change. Attention is divided between four main problems; how previously formed cognitions are organized, what factors affect the reception of new information, the way that different pieces of information are combined to make judgements of an object, and the factors that affect the language used to communicate these decisions to others. The book is divided into three main sections; the first describes the author's approach to the topic, defines the terms used, and briefly discusses problems of measurement; the second is concerned with cognitive organization, that is to say, the way that attitudes and beliefs already held relate to each other; the third discusses the reception and acceptance of new information, and its integration with previous attitudes and cognitions.

It is obvious from this description that the book isn't about cognitive organization, nor about cognitive change in any general sense, but is specifically about attitudes and attitude change, despite the use throughout of 'cognition' and 'attitude' as synonyms; nor is there much in the view adopted by the author to justify the qualifier in the title. From the title, one might expect that a hierarchical model of attitudes was to be presented, perhaps along the lines of Kelly's personal construct theory, with subject specific attitude dimensions; however, not only was this not an important theme in the book, the theories reviewed all fail to specify overall structures for attitudes, and instead leave an uneasy feeling that all our belief systems consist of groups of three attitudes in various states of imbalance.

Disregarding the title, and concentrating instead on the dust cover description of the book as an analysis of research and theory pertaining to the formation and change of beliefs and attitudes, the book does provide a good review of most of the theoretical approaches used to discuss attitude and opinion change, and gives a compact summary of a considerable body of the attitude change literature; Wyer's own approach, which attempts to account for the relations among beliefs and attitudes in terms of simple probability theory using Bayes' theorem, is often shown to provide an acceptable account of data on attitude change that appears paradoxical when viewed in the context of rival theories. However, the extensive literature on subjective probability, much of which violates the simple model proposed, is hardly referred to. Wyer's emphasis on explanations of experimental results in terms of the demand characteristics of the experiment as they appear to the subject also provides a healthy corrective to convoluted explanations in terms of concepts like dissonance often proposed in the past, and highlights a problem common to all psychology experiments, that the subject often conforms to what he thinks the experimenter wants him to do.

Measurement is a crucial issue in this area; a number of quite sophisticated techniques for measuring attitudes have been developed, most of them based on paper and pencil self report techniques. Whilst studies that claim to demonstrate that attitudes measured by such techniques and observed behaviour are inconsistent are not wholly convincing, a fuller discussion of such problems would have been welcome, along with more detailed discussion of some of the methods and implied theory of scaling subjects' responses used in different studies. Some general discussion on the basis for discriminating between theories would also have been useful, in particular on the identifiability of particular theories within the same general class of theory. This is not to say that Wyer is naive in his approach to scaling and measurement – on the contrary, he provides good examples of experimental results that can be as easily accounted for in terms of measurement artifacts as in the theoretical terms proposed by the original authors – merely that the measurement and theory evaluation questions could have been fruitfully discussed in greater detail in the early chapter on rating scales.

Wyer points out that theoretical formulations of social inference phenomena typically oversimplify these phenomena in two ways. First, they implicitly assume that social inferences are governed by a single universal rule, and secondly that they often concentrate only on the evaluative or affective aspects of the information presented and the judgements being made, and ignore their descriptive or denotative aspects. It is unfortunate that he only allows the last chapter (on Abelson's theory of implicational molecules and Gollob's 'subject-verb-object' formulation of social inferences) to deal with alternative formulations which avoid some of these pitfalls. One might suggest that a conjoint measurement approach like the one proposed by Krantz and Tversky (1971) might also be useful in establishing the way that attitudes and beliefs combine to affect judgements of objects, in that hopefully the scaling procedure and the laws of composition relating the attributes to each other can be discovered simultaneously.

The strength of this volume is the comprehensive review and criticism of theory and findings in the area of attitude change; as such it should prove a valuable reference for advanced undergraduate courses and for postgraduate studies. Its weaknesses, such as they are, arise from a lack of a detailed consideration of the measurement problems in attitude research.

JIM RIDGWAY

KRANTZ, D. H. & TVERSKY, A. (1971). Conjoint measurement analysis of composition rules in psychology. *Psychol. rev.* 78, 151–169.

Psychophysics: Introduction to its perceptual, neural, and social prospects. By S. S. STEVENS. Edited by GERALDINE STEVENS. Chichester: John Wiley. 1975. Pp. v+329. £9.50.

S. S. Stevens died in 1973, leaving the manuscript of 'Psychophysics' completed, but not read over. His widow took on the responsibility of editing it and completing the final stages of publication. In such circumstances, one would have been prepared to pardon a few rough passages and ambiguities; in fact, this is not necessary. The book maintains the excellent standard of lucidity that was one of the hallmarks of Stevens' work throughout his life.

The theme of the book is the assessment of sensory scaling methods and laws of psychophysics. Most of the material derives from the work of the author and his students at the Harvard Psychophysical Laboratory (formerly the Psycho-acoustic Laboratory). Nearly all of the experimental results and some of the other material have already appeared in journal articles; it is refreshing to find in a book of this nature that very little reference is made to unpublished data.

The basis of the methodology is that a valid measurement occurs when an observer 'matches' or 'equates' an element from one domain to an element from another domain. 'Domain' normally refers to a sensory continuum, but may also refer to an internalized scale, such as the number scale. In one chapter, *The Scaling of Social Consensus*, tentative extensions are made to other internalized scales such as those underlying judgements of the importance of Swedish monarchs or the desirability of different wrist watches. However the main concern is with sensory psychophysics. Thus, if an observer is presented with a stimulus from a 60 Hz vibrator on his fingertip and asked to match the magnitude of this with that of a band of low frequency acoustic noise, whose intensity he is free to alter, then he is able to respond with a consistent measurement. In fact, subjects are reported to find this case, an example of cross-modal matching, easy and straightforward. On the other hand, a case of acoustic intramodal matching, equating the loudness of a pure tone signal with that of a white noise signal, is found to be much more difficult and this is reflected in the greater variability of the measurements. Stevens contends that this variability can be treated as random fluctuations, and the matching functions that he presents are, in the majority of cases, obtained by combining results from about a dozen observers. If the task consists of matching a sensory continuum to the subject's internalized number continuum, then it becomes the familiar one of magnitude estimation. Stevens now treats this as one special case of cross-modal matching, but argues that this is the purest of numerical scaling methods.

The outcome of matching experiments is the power law, i.e. scaled subjective magnitude is a power function of sensory physical magnitude. Although patently committed to his paradigm and his power law, Stevens is aware of a significant number of unbelievers among the community of psychologists. He gives generous consideration at various points to objections raised and furthermore devotes his final chapter to *Hazards and Remedies*. In 1968 Poulton commented that about one third of the variance found in power law exponents could be accounted for in terms of the range of intensities used in the stimulus continuum. Stevens refutes this in an original way by re-examining the data and discovering that this estimate is too low (the correlation between range chosen and exponent over 21 experiments is −0·94). Stevens' interpretation of this is that the range of stimuli available (below the 'objectionably strong' maximum normally used in matching experiments) is determined by the nature of the sensory transducer. Perhaps, but it could still be that the subject's judgement is based cognitively on his knowledge of the range of stimuli liable to be encountered, using either his memory of stimuli previously presented or his knowledge of sensory ecology. It is not obvious how to separate these viewpoints although some suggestive data on partially hearing patients shows that conductive hearing loss leads to a power function with the same exponent as normals (these patients use a reduced range in matching), whereas neural hearing loss leads to a higher exponent. More work of this nature might be revealing.

Belief that the power law reflects directly the physiological operation of sensory transducers is expanded in a chapter on *Neural Correlates* which is probably the weakest section of the book. It consists of a survey of studies of sense organs in which a power law relationship occurs between neural firing and the physical stimulus. Several cases can be found, particularly if data of the type illustrated in Fig. 83 (Borg *et al.* on lingual nerve responses to sucrose) are held to 'accord with the power law'. However sensory physiologists are unlikely to be satisfied with, for example, a discussion of apparent brightness that has no mention of adaptation mechanisms.

A chapter on *Thresholds and the Neural Quantum* fits rather uneasily with the rest of the book. Signal detection theory is here dismissed in two pages as being an approach whose concern is more methodological than substantive. This is perhaps rather sweeping. The fluctuations in the Békésy audiogram discussed immediately prior to this call to mind response criterion effects – as indeed do other results in the book.

The lasting contribution of Stevens' work may well turn out to be the immense range of sensory phenomena that he cajoled into the laboratory in quantitative form. Those who are not ardent power law enthusiasts will still discover much of interest in this book.

JOHN FINDLAY

An Introduction to Perception. By IRVIN ROCK. New York: Macmillan. 1975. Pp. xii + 580. Price not given.

In a field which lacks any universally-accepted paradigm it is difficult for the would-be author of a textbook to know how best to proceed. 'How does one write a textbook in a field such as perception where there is so little agreement among experts?' As Professor Rock sees it, in the opening paragraph of his Preface, there are two equally unattractive possibilities open to him: either to select for discussion those phenomena which he feels to be important, according to his personal bias, or to attempt to give weight to all the known facts, hypotheses and theories. Not wishing to produce either a dull catalogue or a purely personal statement, he searches for a more acceptable alternative. By the beginning of the second paragraph he thinks he may have found one: for even though there is no generally accepted body of theory, there is 'fairly wide agreement' about the 'central phenomena of perception' and 'what the central problems are'; there is even 'good agreement about the kinds of solutions to these problems that must prove correct'. So he decides to concentrate on the problems arising from these 'central phenomena'.

What, then, are the central phenomena and problems of perception? The author does not provide a direct answer but says, 'for example, the perceptual constancies, illusions of motion, and the like'. However, he does specify the particular chapters (eight in all) which, in his view, fulfil the requirements of a textbook. They cover the perception of size, depth, movement, form and 'neutral color', with particular emphasis being given to problems of perceptual organization, the perceptual constancies, illusions, after-effects, and the nature–nurture issue. It is evident from this list (to which the author has added three chapters of a more 'idiosyncratic' nature, on the perception of direction, 'phenomenal' shape and orientation) that the book is concerned with the 'classical' problems of how we perceive the world of objects and events.

The general plan is to consider, one by one, a number of basic ways in which objects may differ in their appearance (in size, shape, colour, location in space, etc.) and then attempt to arrive at some general principles of perception. Each chapter is introduced by a brief discussion of the question, 'Why is the perception of x considered a problem worthy of attention by psychologists?' But although this might help the reader to appreciate the significance of a particular problem, and is to be welcomed on that account, it is questionable whether a series of such brief and specific introductions, supplemented by the 'brief historical perspective' in the general Introduction, would provide the beginning student with 'the necessary historical and theoretical background', as promised in the publisher's blurb. It seems unlikely that he would gain any clear understanding of how the so-called 'central' problems came to be generated within psychology and formulated in the particular way in which they are presented in this book.

While each chapter provides a very readable and well-documented introduction to the detailed experimental findings related to a particular problem and the main kinds of solution which have been proposed, here again the discussion tends to remain unduly close to the 'concrete facts' of a specific area of research. Within each chapter there is rather more continuity in the argument than one usually finds in a textbook, but there is relatively little between chapters, and the links between the general 'principles' of the final chapter and the conclusions reached in the summaries of earlier chapters are not made explicit. Having worked one's way through some very long chapters dealing with particular aspects of visual perception (e.g. 51 pages on 'size', 56 on 'neutral, color', 67 on 'illusions'), it is also disappointing to find, at the end of it all, only nine pages devoted to 'Principles of Perception'.

This is not to imply that Professor Rock fails to provide any general theoretical discussion elsewhere in the book; he certainly does touch on the wider issues in the Introduction and in several other chapters. But because of the restrictions he has imposed upon himself in his attempt to write an introductory 'textbook', discussion of a more general nature is never allowed to develop very far. This is unfortunate for the beginning student surely wants to be introduced to some general ideas about the nature of perception and not merely to 'concrete facts' and hypotheses about particular phenomena of visual perception. In a subject which is still in the exploratory stage and still characterized by sharply-contrasting approaches, it seems desirable that he should be exposed as soon as possible to the main points of view in the field, if only to help him understand why the accounts he reads of the subject in different textbooks often vary so widely.

In this particular instance, the student needs to know that *An Introduction to Perception* is the kind of introduction you would expect to get from an author who has been brought up within (what Hochberg has called) the 'modern classical tradition', and has retained what some sophisticated contemporary perceptionists might regard as a rather naive and outmoded 'phenomenological' point of view. More specifically, it is the kind of introduction you would expect a 'consecutive' theorist of Helmholtzian ('unconscious inference') lineage to write, for the author believes that percepts are created by a 'rapid-fire, unconscious, process of construction or problem-solving' from ambiguous sensory data. But it is decidedly not the kind of introduction you would expect from the pen of a 'registration' theorist, such as J. J. Gibson, who holds a *radically-different* point of view about the nature of perception, and for whom an ecologically-oriented taxonomy of topics would seem more 'natural' as a starting-point. Nor is it the kind of introduction you would expect from a member of the influential 'feature-analytic' school of perceptual theorists; from a behaviourist (modern or old-fashioned); from an 'information-processing' enthusiast; and so on....

As a readable and well-documented introduction to some of the problems of visual perception, this book should prove useful to second- and third-year psychology undergraduates for the 'visual' section of their course in perception. But it should be regarded as illustrating one particular approach to experimental problems in that field rather than as an authoritative guide to the prevailing ideas in the field as a whole. *An Introduction to Perception* would have served as a better guide if the author had not attempted to write in 'textbook' style, as if the study of perception were a 'normal science', in Kuhn's sense, when it clearly is not.

G. W. GRANGER

The Perception of Light and Colour. By C. A. PADGHAM and J. E. SAUNDERS. London: G. Bell and Sons. 1975. Pp. ix+192 (10 plates). £6.50.

We already have an ample choice of introductory accounts of visual perception, but this new text, written with the bias of physicists, covers areas normally neglected and is therefore to be welcomed. Padgham and Saunders begin with a brief and conventional summary of the optics of the eye and of the neuroanatomy and neurophysiology of the visual system. They include the obligatory and uncritical description of the work of Hubel and Wiesel. Thereafter they concentrate on the perception of light and colour. Colorimetry, colour atlases, chromaticity diagrams, after-images and colour deficiency are covered in particular detail and there are purely physical accounts of colour television, colour printing and colour photography.

The authors have an admirable capacity to explain difficult concepts succinctly and clearly. Only the sometimes casual use of psychophysical expressions and the occasional mixing of discourses will offend the psychological reader: 'Sensation is the fundamental element of a perception...perception (q.v.) results when the sensation is interpreted by the brain.'

Padgham and Saunders are keen that the reader should try simple experiments for himself. One especially happy example is their suggested measurement of dark adaptation using only a star atlas and a list of stellar magnitudes. Since the scale of stellar magnitudes is a logarithmic one, it is possible, on a moonless night, to obtain a very satisfactory dark-adaptation curve by plotting on the ordinate the magnitude of the faintest star visible in a chosen region of the heavens. Although it is often stated that one can see stars down to a magnitude of 6·0, it is often not made clear that this is so only after one has been in the dark for about 25 minutes. Padgham and Saunders add that normal starlight significantly adapts the eye and report that stars of magnitude as low as 7 and 8 are visible if the observer sits in a dark room that has only a very small opening to the night sky.

Most of the text is as accurate as it is clear, but one error is serious enough to deserve note. The authors suggest that strong evidence for the trichromatic theory of colour vision is provided by the microspectrophotometric measurements of single primate cones reported more than a decade ago by Marks, Dobelle and MacNichol and by Brown and Wald; and once more these results are reproduced in a textbook. These much-vaunted, but fragmentary and preliminary, measurements have never been extended and from the start were too noisy and too few to prove the trichromatic theory; indeed if one plotted all the densitometric results in a single figure (cf. Riggs, 1967, *Investigative Ophthalmology* 6, 6–17, Fig. 3) one might rather consider that there was good evidence for a polychromatic theory. What the microspectrophotometric measurements do suggest is that the different pigments underlying colour vision are segregated in different cones.

Padgham and Saunders express the hope that their book will be useful to the general reader, to those working in bordering disciplines, to first-year undergraduates and to 'the sixth-form student who wishes to widen his horizons'. To instruct undergraduates I should still turn to Cornsweet's superb 'Visual Perception' and to excite the interest of the sixth-former I should still use Gregory's dated but ever delightful 'Eye and Brain'; I should certainly press the present work on the graduate student in human performance who was about to put Munsell chips in his tachistoscope.

J. D. MOLLON

Differences in Visual Perception. By JULES B. DAVIDOFF. London: Crosby, Lockwood, Staples. 1975. Pp. v+231. £10.00.

This book begins by challenging the beliefs which we, as naïve realists, are supposed to hold about the nature of perception. Whereas 'commonsense' tells us that we all have the same perceptions when viewing a particular scene, scientific investigation reveals many instances where our perceptions differ. What the author sets out to do is to 'document' a large number of circumstances in which differences in perception arise from the same external input.

An obvious example is the differential effect of light- and dark-adaptation on visual sensitivity: someone who has been in a dark room for half an hour is able to detect a faint light which cannot be seen by another person who has just come in from bright daylight. Many other examples can be found in the extensive literature on the subject and the author has necessarily, though reluctantly, had to be selective. His selection covers a wide range of topics: after-images, adaptation and after-effects, the visual effects of eyestrain and drugs, as examples of differences of a temporary nature; perceptual defects of organic origin, disorders of colour vision, cross-cultural differences and 'attentional processes' (as examples of 'long-term differences'); developmental changes in visual acuity, colour and space perception, visual illusions, eidetic imagery and 'attentional processes'; differences in perception of ambiguous figures; and different ways of interpreting the visual world ('set', 'projective tests', 'ways of seeing'). The complete list of items falling under these, and other, headings is very long indeed and well before they reach the end of it many readers will be prepared to admit that 'commonsense' has misled them and that there really are 'differences in perception'.

If the author's intention were merely to give our commonsense beliefs (as represented in textbooks) a bit of a shake-up, one could conclude that he has succeeded in his aim. But he clearly intended to do more than this for he says that he has attempted to 'document' these differences and that he would like to be able to produce a catalogue of the 'necessary facts'. Unfortunately, the status of many of his 'facts' remains uncertain owing to inadequate documentation and lack of critical comment on experimental conditions and controls. For instance, after stating (in very general, and unnecessarily vague, terms) that after-images 'seem to be connected to a bleaching of the retinal pigmentation', he mentions 'complexities' arising from work carried out in the USSR in the 1950s (no references) in which the investigators 'claim' that it is possible to obtain after-images as conditioned responses to auditory stimuli: 'It is said that eventually the tone produced the after-image....' But on what evidence? Surely the reader ought to be offered something more substantial than the author's uncritical précis of the findings of a single unpublished experiment, in such a controversial area of research? This is not an isolated example: there are many others, and one can only conclude that the author has failed to 'document' (at least in any scientifically-useful sense of the word) the various circumstances in which differences in perception occur. Moreover, as the book claims to deal with the 'extent' of individual differences, it is regrettable that the author has made no use of existing quantitative data on variation between and within individuals. It is also a pity that so many obscurities and errors have been allowed to remain in the text (e.g. the statement (p. 3) that colour vision is a property of the fovea only, not of the extra-foveal retina).

But even if the documentation had been free of such blemishes, the reader would still be left wondering about the ultimate purpose of the ideal catalogue of differences which the author seems so anxious to compile. Why would it be 'nice' to be able to present a catalogue of the 'necessary facts'? Necessary for what? The author never explains but merely throws out a hint in his 'Final Word' (p. 199): 'It should be possible from these findings to theorise on the nature of the processes that cause these differences.' He then goes on to say, 'The reader may have been disappointed that there has been no explicit theorising on these lines.' This reader was bitterly disappointed, not only by the author's reluctance to indulge in 'explicit theorising' about the field as a whole, but also by the lack of theoretical discussion of particular areas of research. Lacking the support of any general theoretical framework and without any clear indication being given of the particular hypotheses they were designed to test, or might conceivably be used to test, the 'findings' appear as bare, isolated (and often shaky), facts. The author attempts no final summing-up about 'differences in visual perception'. One obvious conclusion is that this book is poor value for money.

G. W. GRANGER

Rational Behavior, An Explanation Of Behavior That is Especially Human. By MYLES I. FRIEDMAN. Columbia, South Carolina: University of South Carolina Press. 1975. Pp. xiv+188. $9.95.

Friedman says that an individual is behaving rationally 'when he conceives predictions about observable events and tests them through observations' (p. 5). Moreover, he sees the desire to predict one's environment correctly as the 'fundamental overriding and pervasive motive that gives impetus and direction to behavior' (p. 42). In *Rational Behavior*, he expands these two theses in the light of a picture of how the brain (or mind) works.

On the first thesis, it is not clear how much of what a historian, or artist or book reviewer does can be analysed in terms of conceiving and testing predictions about observable events, yet these are rational activities. What is clear is that Friedman takes a wide view of prediction: a child en route for school predicts the road will take him there (p. 45), a baby crying for food is predicting it will come (p. 43), individuals are interested in making predictions about anything in transition (p. 78). However, even a cavalier attitude to language hardly saves the second thesis, which implies that 'the maintenance of predictability is valued above all else' (p. 114) (which seems to make a completely pleasant surprise impossible), and that negative emotion is 'associated primarily with excessive arousal and with the inability either to predict its reduction or the prevention of its reduction' (p. 117). But, the talk of predictability here is doing no work: it is not the inability to predict as such which is causing the negative emotion, but the inability to alter or avert the annoying event. Describing *this* inability as an inability to predict is merely otiose. Similarly, Friedman analyses insanity in terms of misdirected attempts at prediction, and therapy in terms of correcting them, but this overlooks the possibility of cases where patients may be not so much deluded by false predictions as depressed by true ones.

On Friedman's account the mind (or brain) functions essentially passively, despite the introduction of complicated programming and feedback routines. Behaviour is governed by the matching of current events with neurological traces of past events and programmes based on what happened in the past. Discoveries emerge when events of new types are fitted into previously learned schemes of classification, but there are no accounts either of theorising beyond the setting up of taxonomies, or of the adoption of new schemes of classification – disappointing in an explanation of especially human behaviour. Language plays a surprisingly small part in this account of rationality. It primarily aids communication and the fixing in the mind of fleeting desires (p. 159); Friedman does not see that apart from language there is no sense in speaking of the possession of any but the most elementary concepts, or indeed of the giving or having of reasons – a central element in rationality as normally understood, but not one discussed here. Finally, there is general confusion about the logic of psychological expressions: we are told that automatic thought is involved in the execution of reflexes, that neurological research may produce evidence for intentional behaviour, and that a totally unfamiliar event must be memorised before it can be recognised. Importing of familiar expressions into unfamiliar contexts in this way gives only an impression of clarity. ANTHONY O'HEAR

Br. J. Psychol. (1976), 67, 1, *pp.* 133–136
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not preclude later review)

Arnold, D. B. & Doyle, K. O. *Education/Psychology Journals: A Scholar's Guide.* Folkestone: Scarecrow Press, Bailey Bros. and Swinfen. 1975. Pp. x + 143. £4.20.

Bauman, G. & Grunes, R. *Psychiatric Rehabilitation in the Ghetto.* London: Lexington Books. 1975. Pp. xx + 177. £7.50.

Becker, E. *Angel in Armour.* London: Collier Macmillan, 1975. Pp. xi + 195. £1.95.

Bellack, L. *Overload: The new Human Condition.* New York: Human Sciences Press. 1975. Pp. 223. $9.95.

Bennett, W. L. *The Political Mind and the Political Environment.* London: Lexington Books. 1975. Pp. xxiv + 207. £7.50.

Blair, G. M., Jones, R. S. & Simpson, R. H. *Educational Psychology* (fourth edition). London: Collier Macmillan. 1975. Pp. xiv + 624. £5.50.

Bloch, M. (ed.) *Marxist Analyses and Social Anthropology.* London: Malaby Press. 1975. Pp. xv + 240. £6.75.

Blundell, J. *Physiological Psychology.* London: Methuen. 1975. Pp. 144. 75p.

Broman, S. H., Nichols, P. L. & Kennedy, W. A. *Preschool I.Q. Prenatal and Early Developmental Correlates.* Chichester: Wiley. 1975. Pp. x + 326. £9.75.

Brown, J., Cunningham, W. J., Roberts, J. & Williams, P. *Opening the Door.* London: Routledge and Kegan Paul. 1975. Pp. xviii + 260. £6.95.

Brown, R. & Herrnstein, R. J. *Psychology.* London: Methuen. 1965. Pp. xvi + 762. £12.00; paper, £6.50.

Brunk, J. W. *Child and Adolescent Development.* Chichester: Wiley. 1975. Pp. xv + 464. £4.25.

Bryant, P. *Perception and Understanding Young Children.* Andover: Methuen. 1974. Pp. viii + 189. £1.95.

Buijker, H. C. J. & Rijswijk, M. J. *Trilingual Psychological Dictionary.* vol. 1. English/French/German. Bern: Hans Huber. 1975. Pp. xiv + 355. Fr. 58/DM55.

Burt, C. *The Gifted Child.* London: Hodder and Stoughton. 1975. Pp. vi + 214. £4.85; paper, £2.60.

Busse, T. V., Mansfield, R. S. & Messinger, L. J. *Activities in Child and Adolescent Development.* New York: Harper and Row. 1975. Pp. xiii + 165. $4.95.

Cantwell, D. P. (ed.). *The Hyperactive Child. Diagnosis, Management Current Research.* 1975. Chichester: Wiley. Pp. xii + 209. £7.15.

Centers, R. *Sexual Attraction and Love.* Springfield, Ill.: Thomas. 1975. Pp. xii + 326. $12.95.

Chapanis, A. (ed.). *Ethnic Variables in Human Factors Engineering.* London: Johns Hopkins University Press. 1975. Pp. xii + 290. £9.60.

Christensen, A. L. *Luria's Neuropsychological Investigation.* Copenhagen: Munksgaard. 1975. Text, Pp. 201, £6.90. Manual, Pp. 53, £2.75. Test Cards, 158, £10.00.

Clark, J. M. *Time Out?* London: RCN and NCNUK. 1975. Pp. 68. £1.35.

Cooper, D. E. *Knowledge of Language.* London: Prism Press. 1975. Pp. viii + 196. £1.95.

Craig, R. C., Mehrens, W. H. & Clarizio, H. F. *Contemporary Educational Psychology.* 1975. Chichester: Wiley. Pp. xii + 558. £4.80.

Cull, J. G. & Hardy, R. E. *Counselling Strategies with Special Populations.* Springfield, Ill.: Thomas. 1975. Pp. xiv + 341. $16.95.

Dale, D. M. C. *Language Development in Deaf and Partially Hearing Children.* Springfield, Ill.: Thomas. 1975. Pp. vii + 259. $14.75.

Dally, P. & Harrington, H. *Psychology and Psychiatry for Nurses.* London: The English Universities Press. 1975. £3.95; paper, £1.75.

Davies, D. R. & Shackleton, V. J. *Psychology and Work.* London: Methuen, 1975. Pp. 144. 44p.

de Vito, R. A. & Tapley, R. P. (eds.). *A View into a Modern, State-Operated Mental Health Facility.* Springfield, Ill.: Thomas. Pp. xx + 283. $19.75.

Eliasson, S. G., Prensky, A. L. & Hardin, W. B. *Neurological Pathophysiology.* London: Oxford University Press. 1975. Pp. xiii + 397. £3.00.

Fox, R. (ed.). *Biosocial Anthropology.* London: Malaby Press. 1975. Pp. xii + 169. £6.50.

Freeman, J. *In and Out of School.* London: Methuen. 1975. Pp. 144. 65p.

Gardiner, J. M. & Kaminska, Z. *First Experiments in Psychology.* London: Methuen. 1975. Pp. 144. 65p.

Gold, P. *Please Don't Say Hellow.* New York: Human Sciences Press. 1975. Pp. 54. $6.95.

Gopi, K. *The Awakening of Kundalini.* New York: Dutton. 1975. Pp. xii + 129. $3.25.

Green, C. & McCreery, C. *Apparitions.* London: Hamish Hamilton. 1975. Pp. x + 218. £3.75.

Gregg, V. *Human Memory.* London: Methuen. 1975. Pp. 144. 75p.

Grinder, R. E. *Studies in Adolescence* (third edition). London: Collier Macmillan. 1975. Pp. xviii + 576. £4.50.

Hampstead Child-Therapy Course and Clinic. *Proceedings of the 20th Anniversary Celebrations. Studies in Child Psychoanalysis: Pure and Applied.* London: Yale University Press. 1975. Pp. xii + 175. Price not given.

Holland, N. N. *Five Readers Reading.* London: Yale University Press. 1975. Pp. xvi + 418. £7.50.

Holt, J. *Escape from Childhood: the Needs and Rights of Children.* Harmondsworth: Penguin Books. 1975. Pp. 218. 60p.

Hunt, S. & Hilton, J. *Individual Development and Social Experience.* London: Allen and Unwin. 1975. Pp. 288. £5.95; paper, £3.25.

Kalverboer, A. F. *A Neurobehavioural Study in Pre-School Children.* London: Heinemann Medical Books. 1975. Pp. ix + 134. £4.50.

Levi, L. & Andersson, N. L. *Psychosocial Stress: Population Environment and Quality of Life.* Chichester: Wiley. 1975. Pp. xv + 142. £5.60.

Levitt, R. A. *Psychopharmacology: A Biological Approach.* Chichester: Wiley. 1975. Pp. x + 502. £9.45.

Light, P. K. *Let the Children Speak.* 1975. London: Lexington Books. Pp. xiii + 95. £5.50.

Lindgren, H. C. & Byrne, D. Psychology: *An Introduction to a Behavioural Science* (fourth edition). Chichester: Wiley. 1975. Pp. xiii + 636. £4.80.

Lucas, N. B. C. *An Experience of Teaching.* London: 1975. Weidenfeld and Nicholson. Pp. 196. £4.00.

Lumsden, D. B. & Sherron, R. H. (eds.). *Experimental Studies in Adult Learning and Memory.* Chichester: Wiley. 1975. Pp. xv + 208. £8.10.

McCary, J. L. *Freedom and Growth in Marriage.* Chichester: Wiley. 1975. Pp. 700. Paper, £4.75.

McCormick, E. J. & Tiffin, J. *Industrial Psychology* (sixth edition). London: Allen and Unwin. 1975. Pp. xii + 625. £8.50; paper, £6.00.

Maccoby, E. E. & Jacklin, C. N. *The Psychology of Sex Differences.* London: Oxford University Press. 1975. Pp. xiii + 634. £7.00.

McGurk, H. *Growth and Changing.* London: Methuen. 1975. Pp. 144. 75p.

McLennan, R. *Cases in Organizational Behaviour.* London: Allen and Unwin. 1975. Pp. 282. £8.95.

Mahoney, M. J. *Cognition and Behavioural Modification.* Chichester: Wiley. 1975. Pp. xv + 351. £7.10.

Malamud, N. and Hirano, A. (eds.). *Atlas of Neuropathology* (second revised edition). London: University of California Press. 1975. Pp. xvi + 462. £23.35.

Manocchio, T. & Petit, W. *Families under Stress.* London: Routledge and Kegan Paul. 1975. Pp. vi + 208. £4.95.

Matteson, D. R. *Adolescence Today.* London: Dorsey Press. 1975. Pp. xvi + 400. £3.80.

Martindale, C. *Romantic Progression. The Psychology of Literary History.* Chichester: Wiley. 1975. Pp. ix + 225. £9.45.

Mazlish, B. *James and John Stuart Mill, Father and Son in the Nineteenth Century.* London: Hutchinson. 1975. Pp. xii + 484. £6.50.

Miller, S. *Experimental Design and Statistics.* London: Methuen. 1975. Pp. 142. 65p.

Mitchell, J. J. *The Adolescent Predicament.* Toronto: Holt, Rinehart and Winston. 1975. Pp. vii + 263. $4.25.

Morris, K. T. & Cinnamon, K. M. *A Handbook of Non-Verbal Group Exercises.* Springfield, Ill.: Thomas. 1975. Pp. xxi + 299. $14.95; paper, $10.95.

Mosak, H. H. & Mosak, B. *A Bibliography for Adlerian Psychology.* Chichester: Wiley. 1975. Pp. xiii + 320. £9.45.

Nobbs, J., Hine, R. & Flemming, M. F. *Sociology.* London and Basingstoke: Macmillan Education. 1975. Pp. 354. £4.25; paper, £2.95.

Norman, D. A. & Rumelhart, D. E. *Exploration in Cognition.* San Francisco: W. H. Freeman. 1975. Pp. xvi + 430. $13.50.

Parker, A. *States of Mind: ESP and Altered States of Consciousness.* London: Malaby Press. Pp. 198. £3.95.

Pask, G. *Conversation, Cognition and Learning.* Amsterdam: Elsevier. 1975. Pp. xi + 570. D.fl. 96.00.

Osgood, C. E., May, W. H. & Miron, M. S. *Cross-Cultural Universals of Affective Meaning.* Urbana, Ill.: Universiy of Illinois Press. 1975. Pp. xix + 486. $20.00.

Perlin, S. (ed.) *A Handbook for the Study of Suicide.* London: Oxford University Press. 1975. Pp. xii + 236. £5.75; paper, £3.45.

Peszke, M. A. *Involuntary Treatment of the Mentally Ill. The Problem of Autonomy.* Springfield, Ill.: Thomas. 1975. Pp. xiii + 159. $10.75.

Pugh, D. C., Mansfield, R. & Warner, M. *Research in Organizational Behaviour.* London: Heinemann, 1975. Pp. 114. £1.50.

Reich, W. *Listen, Little Man!* Harmondsworth: Penguin Books. 1975. Pp. 109. 50p.

Reich, W. *The Invasion of Compulsory Sex-Morality.* Harmondsworth: Penguin Books. 1975. Pp. 205. 60p.

Reich, W. *Reich Speaks of Freud.* Harmondsworth: Penguin Books. 1975. Pp. 250. 75p.

Reich, W. *The Mass Psychology of Fascism.* Harmondsworth: Penguin Books. 1975. Pp. 430. £1.00.

Renvoize, S. *Children in Danger.* Harmondsworth: Penguin Books. 1975. Pp. 215. 75p.

Rieke, R. D. & Sillars, M. D. *Argumentation and the Decision Making Process.* Chichester: Wiley. 1975. Pp. vii + 312. £4.30.

Sachs, H. L. *Dynamic Personal Adjustment: An Introduction.* New York: Behavioural Publications. 1975. Pp. 374. $9.95.

Sants, J. & Butcher, H. J. (eds.). *Developmental Psychology.* Harmondsworth: Penguin. 1975. Pp. 501. £3.00.

Shotter, J. *Images of Man in Psychological Research.* London: Methuen. 1975. Pp. 144. 75p.

Sonoyak, S. (ed.). *The Psychiatric Nurse as a Family Therapist.* Chichester: Wiley. 1975. Pp. xx + 251. £4.40.

Sokolov, A. N. *Inner Speech and Thought.* New York: Plenum Publishing Corporation. 1975. Pp. ix + 283. $6.95.

Sperber, D. *Rethinking Symbolism.* London: Cambridge University Press. 1975. Pp. xiii + 153. £4.30; paper, £1.90.

Stammers, R. & Patrick, J. *The Psychology of Training.* London: Methuen. 1975. Pp. 144. 75p.

Stein, M. I. *The Physiognomic Cue Test.* New York: Behavioural Publications. 1975. Pp. 30. $1.95.

Stoller, R. J. *The Transsexual Experiment* (volume 2 of *Sex and Gender*). London: Hogarth Press. 1975. Pp. viii + 316. £7.00.

Stott, D. M., Marston, N. C. & Neill, S. J. *Taxonomy of Behavioural Disturbance.* London: University of London Press. 1975. Pp. viii + 184. £4.75.

Swinn, R. M. *Fundamentals of Behaviour Pathology* (second edition). Chichester: Wiley. 1975. Pp. 595. £6.80.

Sundel, M. & Sundel, S. S. *Behaviour Modification in the Human Services.* Chichester: Wiley. 1975. Pp. xii + 283. £4.95; paper, £2.95.

Ullman, L. P. & Krasner, L. *A Psychological Approach to Abnormal Behaviour* (second edition). Englewood Cliffs, New Jersey: Prentice Hall. 1975. Pp. xxi + 776. Price not given.

Underwood, B. J. & Shaughnessy, J. J. *Experimentation in Psychology.* Chichester: Wiley. 1975. Pp. 236. £5.30.

Uttal, W. R. *Cellular Neurophysiology and Integration.* Chichester: Wiley. 1975. Pp. xviii + 310. £8.00.

Uttal, W. R. *An Autocorrelation Theory of Form Detection.* Chichester: Wiley. 1975. Pp. ix + 158. £7.40.

von Maltitz, F. *Living and Learning in Two Languages.* Maidenhead: McGraw-Hill. 1975. Pp. xiii + 221. £4.85.

Wedell, K. (ed.). *Orientations in Special Education.* Chichester: Wiley. 1975. Pp. x + 210. £7.50.

Weiner, E. *Art and Human Emotions.* Springfield, Ill.: Thomas. 1975. Pp. x + 190. $6.75.

Wexley, K. N. & Yuill, G. A. (eds.). *Organizational Behaviour and Industrial Psychology.* London: Oxford University Press. Pp. xi + 641. £6.38.

Willis, R. *The Interpretation of Symbolism.* London: Malaby Press. 1975. Pp. xv + 181. £6.75.

Zusne, L. *Names in the History of Psychology: a Biographical Sourcebook.* Chichester: Wiley. 1975. Pp. xviii + 489. £9.90.

New Books from Academic Press

THE SOCIAL PSYCHOLOGY OF BARGAINING AND NEGOTIATION

Jeffrey R. Rubin and Bert R. Brown

384 pp.
£7.90/$16.50
0.12.601250.4

Bargaining represents a major mechanism of conflict resolution in our society: this book offers perhaps the first comprehensive summary and integration of the research on bargaining, incorporating more than 500 laboratory studies. In reviewing and integrating the massive literature on bargaining, the authors explore in depth the four basic dimensions of bargaining behavior: structural factors, individual differences, interdependence factors, strategic factors. This volume will be of great value to students and professionals in social psychology, conflict resolution, industrial-labor relations, business administration, international relations, government, political science, law, sociology, and public administration.

Contents

Introduction. The bargaining relationship. An overview of social psychological approaches to the study of bargaining. Social components of bargaining structure. Physical components of bargaining structure. Issue components of bargaining structure. Bargainers as individuals. Interdependence. Social influence and influence strategies. Afterthoughts.

European Monographs in Social Psychology Number 7. Series editor: H. Tajfel

SPEECH STYLE AND SOCIAL EVALUATION

H. Giles and P. F. Powesland

xiv + 218 pp.
£6.60/$17.00
0.12.283750.9

This book presents a thorough investigation into the social importance of individual speech style. The authors examine the results of extensive research and show how listeners make judgements about a person's social status, competence and personality on the basis of his manner of speech. This stimulating study should be of particular interest to sociologists, social psychologists, teachers and students of linguistics, education and speech and communication studies. The lucid presentation of this fascinating subject will also attract a large number of general readers.

COOPERATION: An Experimental Analysis

Gerald Marwell and David R. Schmitt

218 pp.
£7.70/$16.00
0.12.476350.2

Contents

Introduction: the study of cooperation. Methods: measurement and experimental design. The first experiments: inequity and cooperation. Cooperation and interpersonal risk. Cooperation and interpersonal risk: tests for generality: variation in rewards, varying other characteristics of risk, cross-procedural and cross-population replications. Achieving cooperation under risk: protection and communication. Achieving cooperation under risk: relations between people: minimal relations, maximal relations, the invocation of group norms, inequity and interpersonal risk. Achieving cooperation under risk: pacifism. Conclusion and speculations. Appendix: an experiment comparing social coordination in the Cohen–Lindsley and modified settings. References. Author index. Subject index.

Academic Press

London New York San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers
24–28 Oval Road, London NW1, England
111 Fifth Avenue, New York, NY 10003, USA
Australian office
P.O. Box 300, North Ryde, NSW 2113, Australia

EVALUATION OF PSYCHOLOGICAL THERAPIES

Psychotherapies, Behavior Therapies Drug Therapies and Their Interactions

Edited by Robert L. Spitzer, M.D.
and Donald F. Klein, M.D.

Techniques for treating specific psychiatric disorders proliferate, but laws to assure the safety and efficacy of these forms of treatment are not on the books. Attempting to fill the need for objective evaluation of therapies, these papers examine such studies as the Menninger Psychotherapy Research Project.

The contributors and their critics take a searching look at the efficacy and interaction of the most popular psychiatric therapies – individual psychoanalysis and psychotherapy, behavior therapy, and drug-psychological therapies – in treating depression, schizophrenia, behavioral disturbances of children, alcoholism, and various phobias. Based on the proceedings of the 1974 meeting of the American Psychopathological Association, these papers contribute to an understanding of the various psychological treatments and offer the public a chance to get the best return for every therapeutic dollar.

336 pages £11.05

Johns Hopkins University Press
2–4 Brook Street, London W1

PSYCHOLOGY IN AFRICA

MALLORY WOBER

In an exhaustive and well-documented report the author, a psychologist who has himself done research in Nigeria and Uganda and who has lectured at Makerere University, draws together the main threads of the research carried out so far, putting the issues in an African perspective but anchoring them firmly within the framework of modern psychological thinking and technique.

Are there any common personality and intellectual characteristics among Africans? How does weaning affect African child development? How have Africans' feelings developed about city life and industrial work? The questions the author considers range from the broad-based to the specific. The challenges which lie ahead for African investigators now moving into the mainstream of the work are also discussed.

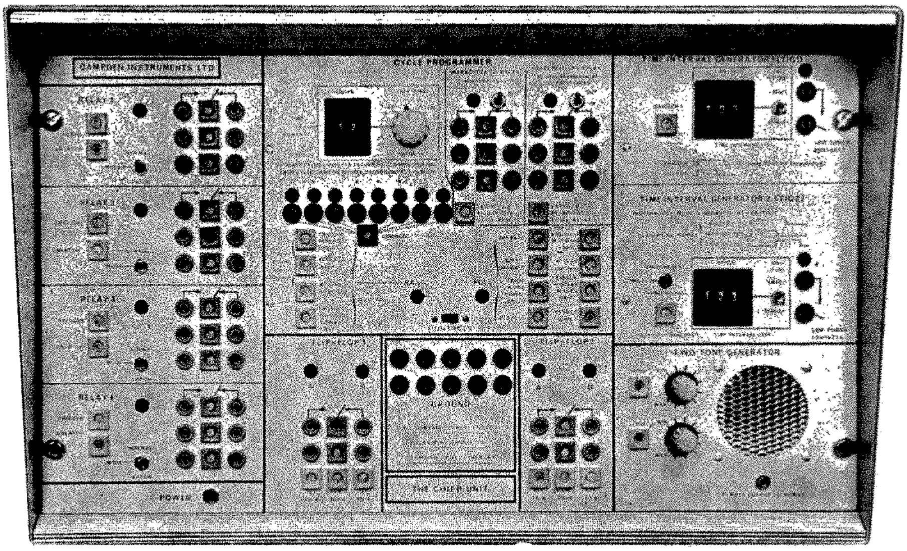
But perhaps above all the book makes a convincing case for psychology becoming a relevant and finely honed discipline in Black Africa, characterised by practical application to Black African society.

220 pp., Figures, References, Index Hardback £4.00 Paper £3.00
0 85302 045 0 Hardback 0 85302 047 7 Paper

Order from your bookseller or from:
THE INTERNATIONAL AFRICAN INSTITUTE,
Room B
210 High Holborn,
London WC1V 7BW
(All orders should be prepaid plus 10 per cent postage and packing.)

IAI members benefit from price reductions on all publications and receive free Africa *and the* International African Bibliography. *Please write for details.*

NEW from CAMPDEN INSTRUMENTS – a powerful and unique package of control logic – THE CHIPP UNIT (Controller for Human Information Processing Psychology)



The CHIPP UNIT has been designed by CAMPDEN in consultation with Dr. John Frisby and Dr. Adrian Simpson of the Department of Psychology in the University of Sheffield who have also written the comprehensive Teaching Manual. The design of the CHIPP UNIT caters ideally for the present day trends in Psychological Experimentation thus superseding existing "Student Laboratory Units"

SPECIAL FEATURES INCLUDE:–

* LOGIC FACILITIES IN A SINGLE INSTRUMENT

* COMPLETE ELECTRICAL ISOLATION BETWEEN THE CHIPP UNIT AND OTHER EQUIPMENT

* L.E.D. INDICATOR LAMPS TO SHOW STATES AT ANY GIVEN INSTANT

* SIMPLE LOGICAL OPERATION BY GROUND SIGNALS

* CLEAR PANEL DESIGN FOR EASE OF OPERATION

* IDEALLY SUITED FOR STUDENT USE AS WELL AS RESEARCH – INCLUDING ANIMAL WORK

The CHIPP UNIT comes with a comprehensive USER'S MANUAL which includes a full description of how it is operated. This Manual includes illustrations on the programming of a number of psychological experiments and is written with the student user particularly in mind so that students should be able to teach themselves how to use it.

For more information on the CHIPP UNIT also on our FISH experimental units write to:–

CAMPDEN INSTRUMENTS LTD.,
186, CAMPDEN HILL ROAD, Telephones:– (01) 727 3437
LONDON, W8 7TH. (01) 229 3636

The British Journal of Criminology

Delinquency and Deviant Social Behaviour

The purpose and policy of the Journal is to foster and publish researches and studies into the causes and prevention of crime, and the treatment of criminals.

Each number of the journal contains original papers. In addition the journal publishes general surveys of current developments in various countries. On occasion an individual part of the journal will consist of a symposium dealing with a special subject, with papers from groups of experts on issues not only purely scientific but of public moment at any given time.

There are in-depth notes and comments on practical developments, notably in the field of penal reform, on government or privately sponsored reports, and on new experiments in dealing with delinquent or deviant persons. Finally the Books and Periodicals section provides a comprehensive review of all significant literature in the field.

MAIN CONTENTS OF THE OCTOBER 1975 ISSUE

Psychiatric Services and the Magistrates' and Juvenile Courts, by Herschel Prins.

Children's Hearings: A Retrospect, by Nigel Bruce.

Children's Hearings: A Commentary, by P. D. Scott.

Suicide and War, by Pat O'Malley.

Self-Labelling Theory: Preliminary Findings Among Mental Patients, by Mordechai Rotenberg.

A Rehabilitation Hostel for Drug Users: One Year's Admissions, by Christopher Melotte.

Bright Delinquents, by Gavin Tennent and Dennis Gath.

Annual Subscription: £9.50 post free.

Published quarterly on behalf of The Institute for the Study and Treatment of Delinquency by

Stevens & Sons Ltd.

Subscriptions: North Way, Andover, Hants, SP10 5BE

British Journal of Educational Psychology

Vol. 46, Part 1

February 1976

The British Journal of Educational Psychology is concerned mainly with reporting the results of experimentally based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry.

The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words).

There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists.

Subscriptions (£6 a year, or £2 per issue) should be sent to: Scottish Academic Press Ltd., 33 Montgomery Street, Edinburgh EH7 5JX.

CONTENTS

Symposium: Learning Processes and Strategies.

Editorial Introduction: The Verb 'to Learn' Takes the Accusative.

Marton, F. and Saljo, R.: On Qualitative Differences in Learning: I. Outcome and Process.

Pask, G.: Conversational Techniques in the Study and Practice of Education.

Articles

Donachy, W.: Parent Participation in Pre-School Education.

Galloway, D.: Size of School, Socio-Economic Hardship, Suspension Rates and Persistent Unjustified Absence from School.

Clifford, M. M.: Are Common Sense Decisions Deterring Learning?

Tamir, P.: The Relationship between Achievement in Biology and Cognitive Preference Styles of High School Students.

Biggs, J. B.: Dimensions of Study Behaviour: Another Look at A.T.I.

Research Notes

Stanley, G.: The Processing of Digits by Children with Specific Reading Disability (Dyslexia).

Shayer, M.: The Pendulum Problem.

Weir, C.: Interviews as an Aid to Selection of Psychology Undergraduates.

Book Reviews

Publications Received

SCOTTISH ACADEMIC PRESS LIMITED

33 MONTGOMERY STREET, EDINBURGH EH7 5JX

REVUE DE PSYCHOLOGIE APPLIQUEE

Publication Trimestrielle

Directeurs: Dr P. PICHOT et P. RENNES

3ème Trimestre 1975 – Volume 25 – No. 3

SOMMAIRE

J. LALANNE – Une approche de développement moral.

R. PERRON – L'élaboration du récit au D.P.I.
L'organisation de l'univers humain.

KARL H. PRIBRAM – Portrait, notice biographique et bibliographique.

Rédaction et Administration:
48 Avenue Victor Hugo, 75783 PARIS CEDEX 16

Abonnements (un an) France: 45 F; Étranger: 55 F
Le numéro France: 12 F; Étranger: 12 F

Règlements au nom des Éditions du Centre de Psychologie Appliquée

Numéro spécimen sur demande

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA

rivista trimestrale

pubblicata a cura dell'Università Cattolica del Sacro Cuore

direttore Leonardo Ancona *comitato di redazione* Mario Bertini – Renzo Carli – Fabrizio Napolitani – Luigi Pizzamiglio – Corrado Pontalti – Carlo Saraceni – Ciro Sdino

Year XXXVI Nr. II April–June 1975

Leading article.

Zappella, M. Ethological proposal and child's autism.

Prandi, G. Instincts, affection and phantasm.

Cavazzuti, V. The effect of massive and distributed exercise after monolateral temporal lobectomy in man.

Mancia, M. Sleep, dream and instincts: an interdisciplinary approach.

Bassan, F. Psychanalysis and philosophy in Money-Kyrle's concepts. I.

Ancona, L. – Croce, M. A. – Lisi, F. Duration of filmic effects in medium-high learned individuals.

NOTES AND DISCUSSIONS

Giossi Morbelli, L. Perspectives and interpretations of command activity.

Barolo, E. Syncretism in child's perception.

Every article is followed by an Italian, French, English and German summary.

Every issue summarizes many Italian and Foreign works of psychology, neurology and psychiatrics and gives large review of news and information.

Redaction – Administration: 1, Largo A. Gemelli – 20123 MILAN (Italy).

Annual subscription 1975: Italy L. 10.000; foreign countries L. 15.000 (or equivalent foreign money).

CALCUTTA

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
The Journal of Occupational Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *The Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1976 subscription prices, payable in advance, are: £20.00 net per volume post free (US $60.00 in the USA and Canada) for *The British Journal of Psychology*; £16.00 net (US $48.00) for *The British Journal of Medical Psychology*; £13.00 net (US $39.00) for *The Journal of Occupational Psychology*; and £19.00 net (US $57.00) for *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian inquiries should be addressed to the Cambridge University Press, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15.00 per volume post free ($34.00 in the USA). Subscriptions should be sent to the Society's Agents, John Wright & Sons, Ltd., 42–44 Triangle West, Bristol BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £6.00 a year post free, payable in advance, or £2.00 each number. The subscription price in the USA is $20.00. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

The series of *Monograph Supplements* issued in connexion with *The British Journal of Psychology* is now discontinued. Orders may be sent to any bookseller, or the Cambridge University Press, Bentley House, 200 Euston Road, London NW1 2DB, from whom details of these publications may be obtained.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. D. B. Clarke, Department of Psychology, The University, Hull HU6 7RX;

for *The British Journal of Medical Psychology* should be sent to

Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London SW20 0NE;

for *The Journal of Occupational Psychology* should be sent to

Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster LA1 4YF;

for *The British Journal of Social and Clinical Psychology* should be sent to

Dr Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ: or to
Dr H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester M20 8LR;

for *The British Journal of Educational Psychology* should be sent to

Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price 50p net (USA $2.00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18.00 (£16.00 if paid by Banker's Order). They are entitled to purchase any or all of the Society's Journals at special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London W1X 4DN.

The British Journal of Psychology *Volume 67, Part 1, February 1976*

CONTENTS

ROWELL, J. A. AND RENNER, V. J.
Quantity conceptions in university students: another look *page* 1

BENJAFIELD, J. AND ADAMS-WEBBER, J.
The golden section hypothesis 11

NETTELBECK, T. AND LALLY, M.
Inspection time and measured intelligence 17

DERĘGOWSKI, J. B.
Implicit-shape constancy as a factor in pictorial perception 23

MORTON, JOHN AND CHAMBERS, SUSAN M.
Some evidence for 'speech' as an acoustic feature 31

GUPTA, B. S.
Extraversion and reinforcement in verbal operant conditioning 47

EYSENCK, MICHAEL W.
Extraversion, activation and the recall of prose 53

DAVIES, GRAHAM AND PROCTOR, JOHN
The recall of concrete and abstract sentences as a function of interpolated task 63

BOUCHER, JILL AND WARRINGTON, ELIZABETH K.
Memory deficits in early infantile autism: some similarities to the amnesic syndrome 73

CAMPBELL, ROBIN, DONALDSON, MARGARET AND YOUNG, BRIAN
Constraints on classificatory skills in young children 89

WARDEN, DAVID A.
The influence of context on children's use of identifying expressions and references 101

BOOK REVIEWS 113

OTHER PUBLICATIONS RECEIVED 133

CALCUTTA UNIVERSITY

© The British Psychological Society 1976

10 DEC 1976

Printed in Great Britain at the University Printing House, Cambridge

Volume 67, Part 2 *May 1976*

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by A. D. B. CLARKE
with J. BROWN, R. CONRAD, N. F. DIXON,
G. W. GRANGER, I. MARTIN, R. L. REID,
B. TIZARD *and* P. B. WARR

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

CAMBRIDGE UNIVERSITY PRESS

BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB

32 EAST 57TH STREET NEW YORK, N.Y. 10022

Assistant Editor: D. I. WILLIAMS
Book Review Editor: D. C. KENDRICK
Press Editor: MICHELE BENJAMIN

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements.

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of articles published from Vol. 56 (1965). A brief summary, suitable for use as an abstract should precede the article.

Manuscripts should be submitted in duplicate

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the author's names and corresponding to an alphabetical list at the end of the paper. This list must give author's names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–284.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Anée psychol.* **50**, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

© The British Psychological Society 1976

Permissions

For permission to reproduce material from *The British Journal of Psychology*, please apply to the London or New York office of Cambridge University Press.

ISI Tear Service, 325 Chestnut Street, Philadelphia, Pennsylvania 19106, USA is authorized to supply single copies of separate articles for private use only.

Subscriptions

The British Journal of Psychology is published quarterly.

The subscription price (including postage) of volume 67, 1976, is £20.00 net (US $60.00 in the USA and Canada). Single parts cost £6.00 each (US $18.00 in the USA and Canada).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1976), 67, 2, *pp.* 137–147
Printed in Great Britain

ACQUIRED SENSORY CONTROL OF SATIATION IN MAN

By D. A. BOOTH, M. LEE and CHRISTINE McALEAVEY
Department of Psychology, University of Birmingham

Lean male and female subjects who had been taking their meals as usual were given a 100 ml drink of starch immediately before a sandwich lunch. In the training phase, 65 per cent starch was given before lunches which included a yoghurt-based dessert of one flavour and 5 per cent starch was given before another flavour. Subjects who initially ate lunches of similar size following the two drinks began after several pairings to take larger lunches following dilute starch than following concentrated starch. In extinction tests, identical 35 per cent starch drinks were given before lunches including desserts of either flavour. Lunches including the flavour hitherto paired with dilute starch remained larger than lunches including the other flavour, at least initially in extinction. Some of the difference in intake occurred in the dessert itself. None was attributable to differences in the early stage of the lunches. These results extend to man recent demonstrations that the satisfying power of a foodstuff is in part acquired by association of its sensory characteristics with some consequence of ingesting the nutrients which are consistently taken with or in that food. This consequence may be rapid intestinal absorption of glucose during or shortly after sensory input. The nature of the intake-suppressing response remains to be determined.

Accounts of the factors controlling feeding behaviour in men and animals consuming meals comprised of a variety of foodstuffs are at present loose and incomplete, based on little direct experimentation, and certainly not quantitative. We are still at the stage of qualitative analysis – establishing relevant phenomena experimentally.

It has been claimed that nutrient uptake can be controlled in the absence of stimuli from the mouth and throat, because animals (Teitelbaum & Epstein, 1963) and men (Jordan, Stellar & Duggan, 1968) stay in fair health if a manipulandum which delivers a complete nutrient directly to the stomach is made available to them. Successful self-infusion depends, however, on maintenance of the rat's activations of the manipulandum by the occurrence of oropharyngeal cues from the infusion or correlated events at times when the animal is close to the manipulandum (Holman, 1969). In any case, the dispensability of sensory cues would imply nothing as to their importance when they are provided as in normal feeding. They are likely to be crucial in facilitating ingestion of solids early in life (Booth, Stoloff & Nicholls, 1974) and in adjusting the selection of foodstuffs to meet nutrient needs (Richter, 1942–3; Rozin & Kalat, 1971; Booth, Lovett & McSherry, 1972; Simson & Booth, 1973). It has also been proposed that food and water intake is measured to an appropriate quantity by oral stimuli (Bellows, 1939; Rozin & Kalat, 1971). This notion of oral metering has been widely invoked, under the possibly mistaken impression that gastric distension or the early stages of absorption (Booth & Davis, 1973; Davis & Campbell, 1973) would not be sufficiently strong or accurate to permit the adjustment of the sizes of normal fairly small meals according to their nutrient value. Yet the specific and general satiating effects of prolonged contact with an initially acceptable flavour such as sweetness (Le Magnen, 1955; Miller, 1957; Hsaio & Tuntland, 1971; Holman, 1973), and the appetite-stimulating effect of change from one familiar flavour to another (Le Magnen, 1956*a*, *b*, 1961), seem ill suited to provide general mechanisms for controlling the quantity of food ingested with respect to its nutrient value. A versatile and potentially precise oral metering system would be available if the decay of involuntary interest in a food as it is eaten could be classically conditioned by nutritional effects of the food, or if the choice of how much of a foodstuff to eat could be instrumentally learned by predicting its after-effects from past experience of food of similar sensory qualities. Le Magnen (1956*c*, 1959, 1969) proposed that satiety could be conditioned by consistently pairing a food odour with some manipulation that affected appetite. Acquired olfactory or gustatory control of the inhibitory influences on ingestion towards the end of a meal (as distinct from acquired aversion) was recently demonstrated in the rat (Booth, 1972; Booth & Davis, 1973; Davis & Campbell, 1973). The present experiment was an attempt to demonstrate the same phenomenon in man. It seemed possible that associative nutritional control of the satisfying power of a foodstuff is particularly important in man. Current feeding habits seem likely to restrict the operation of other mechanisms for behavioural regulation of human energy balance: intervals between meals are strongly determined by social conditions and the size of a meal is largely set at its beginning when the plate is filled.

The experimental meals were as normal as practicable, given the variables of interest. Starch was the nutrient found to be most effective at establishing sensory control of satiety in the rat (Booth, 1972), and so the starch content of the meal was varied by starting it with a dilute or concentrated starch load. A portion of such a load rapidly passes into the intestine and is digested and absorbed during the rest of the meal. This procedure therefore provided considerable opportunity for ingested nutrient to produce effects in body tissues as well as in the gut at the time or shortly after distinctive sensory cues incorporated in the meal had been experienced. The flavour of one of the dessert items was varied according to the composition of the load, to provide a predictive stimulus which might come to affect the amount taken during the meal. Development of changes in the satiating power of the loads consequent on adaptation to the training conditions was distinguished from conditioning or learning, by means of an extinction test, i.e. giving meals which included the distinctive dessert flavours but were preceded by starch loads of a single intermediate concentration.

METHOD

General procedure

Subjects

University staff and students in good health and of normal body build were solicited by advertisement to be given sandwich lunches without charge or remuneration, taking a nutrient-containing drink beforehand. The subject's normal meal pattern and experimental menu preferences were ascertained by questionnaire. Subjects were excluded from this experiment if their typical lunch was less than 200 kcal. He (or she) agreed to treat the experimental meal as a normal lunch and to keep to a standard pattern of breakfast and evening meals and physical and social activities on experimental days. A series of appointments was then made for a time of day as close as practicable to the subject's usual luncheon period. Other experiments, on non-acquired factors in control of appetite by nutrient ingestion, were being run in other subjects. The subjects were informed that the experiments were concerned with the effects of nutrients on appetite but were told nothing of the design of the specific experiment.

Loads

When the subject came for lunch, he was given a lemon-flavoured drink. He was told to gulp it down from the back of the throat and then to wash out his mouth, expectorating all the water. This was to minimize the extent to which the drink was tasted. The drink was 100 ml of either lemon squash (on an adaptation day) or one of two experimental loads. The concentrated load contained 65 g maltodextrin (MD 05, a partially hydrolysed starch, low in glucose and maltose content, soluble in cold water, supplied by Manbré Sugars, Hammersmith), 3 g uncooked cornflour (to make the drink cloudy like the dilute load), 0·5 g citric acid, 0·004 g sodium saccharin, and 0·4 g lemon essence (E. F. Langdale, Croydon). The dilute load contained 4·5 g pregelled starch (Instant Clearjel, Laing-National, Manchester; this thickened the drink, to approach the syrupy texture of the concentrated load), 5 ml of double cream (approximately 2 g of triglyceride, to inhibit gastric clearance similarly to concentrated starch), 2 g sucrose, 0·4 g citric acid, 0·01 g sodium saccharin, and 0·4 g lemon essence. The drinks were presented in a heavy china cup, to disguise the difference in density.

Lunch consumption

Immediately after taking the drink, the subject was presented with an array of luncheon items. These were his previously expressed preferences from a menu including sandwiches (ham, cheese or salmon and cucumber), pork pie, yoghurts of a wide variety of flavours, chocolate-covered biscuit (Kit Kat) or confectionery (Mars Bar), and cold lemon drink or coffee with sugar and milk as desired. Each item was presented in a standard portion amounting to a few mouthfuls only, e.g. an eighth-round of sandwich, a quarter-carton of yoghurt, or a single finger of chocolate biscuit. Two such portions of each selected item were presented at any one moment during lunch, another portion being added to the array when the subject took one. No constraint was put on the sequence in which subjects took samples. A record was made of the time at which each portion was taken. The subject was required to stay for 30 minutes, however soon he had finished. The instruction was for the subject to take as much as he wanted and before he left he was asked whether he had had enough to eat. The amounts taken were determined by weighing item stocks and containers before and after the meal. Energy content of the food consumed was calculated from the weight eaten, using conversion values obtained from manufacturers or derived from weighing the components of item portions and applying data for standard foodstuffs given by McCance & Widdowson (1969) or Church & Church (1970).

The freely chosen flavour of yoghurt was provided in adaptation lunches. In experimental lunches, following concentrated or dilute loads, a modified yoghurt was provided. The subject was asked to select two unfamiliar but acceptable flavours – either never eaten before or very seldom experienced. To make this item even less familiar, the texture of the yoghurt was modified by adding 4 ml of 4 per cent pregelled starch and 1 ml of 25 per cent maltodextrin to 5 g of yoghurt.

Acquisition and extinction

Each of the two unfamiliar yoghurt pudding flavours was consistently presented in a lunch following a particular experimental load. This pairing of sensory cue with consequence-inducing load (conditioned stimulus with unconditioned stimulus, or discriminative stimulus with reinforcer) was termed the acquisition phase of the experiment. This continued for at least four pairs of lunches or until two pairs of lunches in succession were smaller following the concentrated load than following the dilute load. An extinction phase ensued, in which pairs of lunches differing in yoghurt pudding flavour continued, but every lunch was preceded by an identical load – a mixture of 50 ml of concentrated load with 50 ml of dilute load.

Debriefing

Recall of the previous experimental day, including any comments on the effect of the experimental lunch, was elicited by questionnaire when the subject had finished eating. Sizes and timing of meals and snacks earlier on the experimental day and the previous evening were also recalled, and the normalcy or otherwise of the day's activities. This was to encourage adherence to the subject's standard pattern for experimental days.

Additional questions were asked on the last day of experimentation for each subject. These inquired after differences in either taste or effect of the experimental drinks on different days. If the subject thought he detected a difference, he was asked to describe the difference and to identify the direction of difference between the two loads. If a difference in effect was described, the subject was asked to guess the cause of those effects.

Group variations

Initial effect of load

Subjects were divided into two groups according to the relation between the composition of the load at the start of the lunch and the weight of food eaten during the meal. In unpublished experiments, we have found that a direct satiating effect of the concentrated starch load operates strongly if there is a delay of 15 min or more between loading and the start of the meal. When the load is given immediately before the meal, however, there is little effect. Indeed, in at least half the subjects there is no sign of a relative reduction in meal size following the concentrated drink – evidence that its nutrient content is well disguised. Where a meal-size response to an immediately prior load difference exists from the start, the elaboration of a differential response under strong sensory control would not be expected. Therefore, subjects taking less food in the first lunch following a concentrated load than in the first lunch following a dilute load were assigned to the group showing an initial effect, which was not expected to show acquired sensory control in the extinction phase of the experiment. When the sizes of the lunches were the same within 10 g on the first pair of experimental days, or the lunch following the concentrated load was larger than that following the dilute load, the subject was assigned to the group showing no initial effect which should be capable of elaborating differences in meal size according to sensory characteristics of the meal.

Replications

Both types of group, with four subjects in a group, were run at the University of Sussex. The presence of an initial satiating effect of the concentrated load was determined from the first pair of experimental lunches. Another pair of four-subject groups was run later at the University of Birmingham. The initial effect in these subjects was determined in a pair of meals prior to the experimental series, using ordinary yoghurts of the subject's most preferred flavour. There were two women and two men in the initial-effect group in Sussex, whereas the no-initial-effect group was all male. At Birmingham, there were two of each sex in each group. The sequence of yoghurt pudding flavours was strict alternation in Sussex, and there was a single pair of lunches in the extinction phase. The Birmingham flavour sequence was a quasi-random alternation, with three lunch-pairs in extinction.

Statistics

All quoted P values were calculated according to two-tailed Wilcoxon matched-pairs signed-ranks tests.

RESULTS

Subjects who showed no initial reduction in the weight of food taken for lunch immediately following the concentrated starch load relative to the weight taken following the dilute load (Fig. 1, upper panel) did begin to show such a reduction with further experience. All eight of these subjects showed this differentiation in meal size before or at the fourth pair of experimental lunches ($P < 0{\cdot}01$); six out of eight showed it on the third day-pair ($P < 0{\cdot}05$). In subjects who showed an initial satiating effect of concentrated load relative to dilute load (Fig. 1, lower panel), the meal-size difference did not systematically increase during the acquisition phase.

On the first pair of extinction days, on which an identical load of intermediate concentration preceded each lunch, the group meal-size differentiation continued unabated in the subjects who had shown no effect in their first pair of meals (Fig. 1)

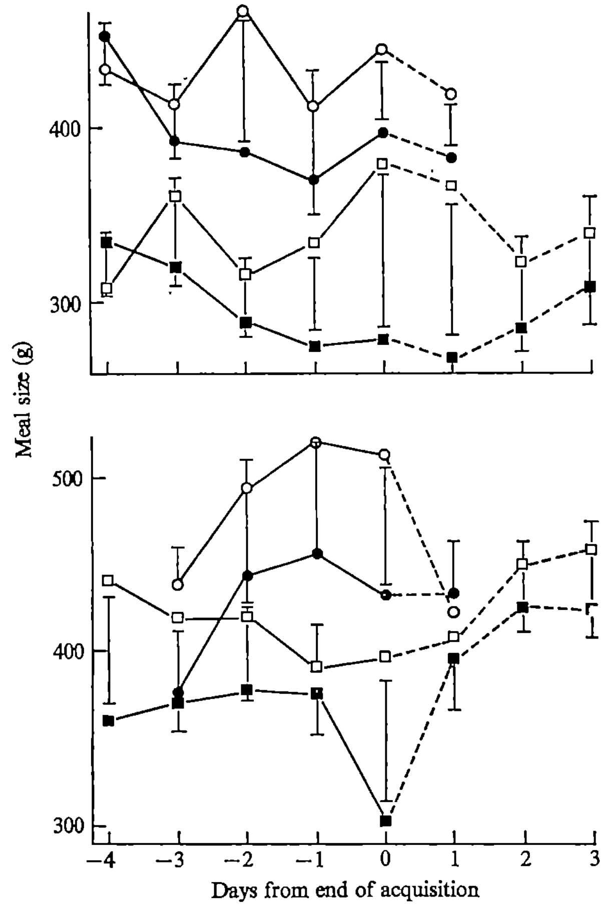


Fig. 1. Acquisition and extinction of meal-size control, in subgroups with no initial effect of load difference on meal size (upper panel) and with an initial effect (lower panel). The acquisition phase (continuous lines) is synchronized on its final day ('backwards learning curve'). The extinction phase (broken lines) was one meal-pair in one pair of subgroups and three meal-pairs in the other pair of subgroups. Filled symbols: meal associated with concentrated starch load. Open symbols: meal associated with dilute load. Vertical bars: S.D. of group difference within the meal-pair. Circles are the data for the Sussex groups and squares are Birmingham data.

and all eight subjects showed the differentiation in that direction ($P < 0{\cdot}01$). The meal-size difference was considerably attenuated on the second pair of extinction days in the group in which they were run, three of the four subjects showing a reduced difference and one no longer showing the effect at all.

In contrast, in the subjects who had shown a satiating effect in their first pair of meals, the meal-size differentiation did not appear at all under extinction conditions (Fig. 1, lower panel).

Despite considerable variation in the caloric density of the items on the lunch menu, the above results for weight of food consumed were paralleled by the energy content of the meals (Table 1). Fluid intakes did not vary systematically between conditions.

Table 1. *Characteristics of experimental lunches*

Variable	Day-pair	Concentrated related meal	Dilute related meal	S.D. of difference	*P*
	Group with no initial meal-size effect				
Total meal size (kcal)	Initial acquisition	835	832	78	n.s.
	Final acquisition	760	946	164	< 0·01
	First extinction	743	948	178	< 0·01
Total meal size (g)	Initial acquisition	391	371	29	< 0·05
	Final acquisition	339	413	72	< 0·01
	First extinction	327	394	63	< 0·01
Size of yoghurt portion (g)	Initial acquisition	139	135	63	n.s.
	Final acquisition	120	174	53	< 0·01
	First extinction	107	161	54	< 0·01
	Group with an initial meal-size effect				
Total meal size (kcal)	Initial acquisition	921	1105	181	< 0·02
	Final acquisition	905	1061	193	< 0·05
	First extinction	983	987	170	n.s.
Total meal size (g)	Initial acquisition	368	432	65	< 0·01
	Final acquisition	367	455	104	< 0·02
	First extinction	415	415	59	n.s.
Size of yoghurt portion (g)	Initial acquisition	120	162	32	< 0·02
	Final acquisition	144	177	50	< 0·02
	First extinction	158	150	42	n.s.

The group means for the amount of food taken in the first 5 min period of lunch were almost identical following drinks containing high or low concentrations of starch, rising from 82·1 g (S.D. 28·8) in the first pair of meals to 98·7 g (S.D. 37·6) in late acquisition and extinction. Mean values for feeding rate in the second 5 min period were 20–35 per cent lower after concentrated starch than after dilute starch throughout the experiment in the group showing an initial meal-size difference, although the difference was never statistically reliable. There was no difference at the start of acquisition in the 5–10 min intakes of the group showing no initial satiation (91·1 g, S.D. 48·0, after concentrated starch; 94·6 g, S.D. 35·4, after dilute starch). However, intake after the concentrated load was lower in that period at the end of acquisition (58·5 g and 83·8 g, $P < 0{\cdot}05$) and at the start of extinction (64·4 g and 91·9 g, $P < 0{\cdot}05$). It was not possible to compare feeding rates in the third 5 min, as some subjects finished lunch in less than 15 min.

About three-quarters of the difference between total amounts eaten in lunches of the two types was attributable under both acquisition and extinction conditions to intake of the yoghurt pudding, whose relatively unfamiliar flavour was consistently associated with a particular load during acquisition (Table 1). The greater intake of yoghurt pudding having the flavour related to the dilute load virtually disappeared in the second pair of extinction days (present in one out of four subjects). A statistically reliable effect was specific to the yoghurt item; differences in the same direction were seen in other items taken late in the meal, and in all the non-yoghurt items taken together, but did not reach statistical significance ($P > 0{\cdot}1$). The yoghurt effects did not appear in the subjects who had shown an initial satiation effect.

At the end of their time in the experiment, nine subjects recalled that there had been two distinctive experimental drinks prior to lunches and seven subjects (five of them in the group showing no initial satiation) recalled three drinks. The words the subjects used to describe how they discriminated indicated that, as anticipated, there was no fully effective disguise for the texture cue which is provided during swallowing by the non-Newtonian fluid properties of a concentrated starch solution (Wood, 1968; Shama & Sherman, 1973). All the group with an initial meal-size effect and no acquired sensory control recalled one of the drinks to have been thick or syrupy and half recalled the other drinks as thin or easy to swallow. The group who acquired sensory control included only five such recalls. Instead they mentioned cues not recalled by the other group, such as smoothness or creaminess (there was cream in the low-starch drink) or difference in taste or flavour (the low-starch drink was slightly sweeter). Thus differences in flavour, appearance and weight attributable to starch content were more successfully masked. The subjects' guesses as to the nutrient correlate of sensory discriminated differences between experimental drinks were correct in six cases from the initially satiated group and three cases from the acquired control group (two-thirds with high confidence in each group). One incorrect guess and one refusal to guess were made in the initially satiated group and two errors and three refusals in the acquired control group. These guesses were mostly expressed using the term 'nutrient' which was used in the question. However, one subject in each group used the word 'sugar' and two subjects in the initially satiated group used the word 'starch' or 'carbohydrate' – all guesses in the correct direction. The subjects in the acquired control group related the sensed creaminess to 'fat' or 'calories', guessing nutrient richness in the incorrect direction. Four of the subjects in the initially satiated group recalled making their nutrient guess early in the experiment and all these guesses were in the correct direction. Two of these four were the only subjects in the experiment who claimed to have noticed consistent differences in their lunch intake which they related to the type of experimental drink taken at the start of the meal and they related them correctly. None of the subjects in the acquired control group recalled having made any nutrient guesses before the end of the experiment.

No subject purported to detect any difference between the two flavoured yoghurt puddings which could have reflected illusory differences in nutrient content or direct satiating power corresponding to the acquired sensory control of intake. On the other hand, two subjects in the initially unsatiated group and four subjects in the initially satiated group guessed (incorrectly) that the yoghurts (like the drinks) varied from the beginning in their intended effect on appetite, and one of the latter four detected the correlation between type of drink and flavour of yoghurt. Nevertheless, there was nothing in any subject's questionnaire responses which indicated an awareness that the intake difference between the two yoghurts was a dependent variable related to a difference between the experimental drinks with which each lunch began.

DISCUSSION

The results clearly show that the total amount eaten during a meal, and the amount eaten of a particular dessert item in the meal, can come under the control of the flavour of that item according to an arbitrary contingency previously imposed between that flavour and the carbohydrate content of an item that hitherto started the meal. The acquired differential sensory control was seen, not only in meals differing in initial carbohydrate concentration, but also just as strongly in meals starting with drinks of identical intermediate concentration. This excluded the possibility that differentiation of intake depended merely on developing adaptation to the initial loads, or an increasing discrimination of their different satiating effects. It also made it unlikely that a compound stimulus of the carbohydrate load's sensory or satiating properties and the dessert's flavour gained control of feeding. The data do not exclude the possibility that some adaptation or learning occurs in the other parts of the meal, which did not vary in flavour and so would have to reflect discriminable differences in preloads. However, if such effects occurred they were smaller than the acquired control by dessert flavour and statistically insignificant under these conditions.

If the different satiating properties of the carbohydrate loads were important, as indeed is the hypothesis, the results suggest that this satiation is an event later in time that the consumption of the flavoured dessert, serving as a consequence for which the dessert flavour can become a cue. This interpretation is consistent with our findings on the time-course of direct satiating effects of concentrated starch loads in man (Booth, Lee & McAleavey, in preparation) and possibly related findings with concentrated glucose loads (Booth, Campbell & Chase, 1970; Cabanac, Pruvost & Fantino, 1973): the strongest satiating effect of such loads generally occurs at 20–40 min after ingestion, by which time most of the subjects in the present experiment had finished their lunch. Dessert flavours would of course be subject to the greatest recency effects in an association with postprandial events.

The nature of the response acquired to the contingency between dessert flavour and effects of carbohydrate ingestion remains to be determined. The present data do not distinguish, for example, between on one hand a learned instrumental adjustment of the amount eaten with the intention of producing an optimum degree of satiation following lunch and on the other hand an involuntary reaction to the flavoured dessert (and perhaps any items which follow it in the meal) as part of a behavioural pattern which expresses an expectation established by the satiation effects of the loads serving as unconditioned stimuli. The energy difference between the carbohydrate loads (about 230 kcal) was approximated by the acquired differentiation in meal size (about 200 kcal: Table 1). However, we do not know the relative satiating effects of carbohydrates and mixed foods or how they would guide an intentional response, neither do we know the relation between the intensity of the unconditioned satiation stimulus and the vigour of the resulting conditioned satiation response. So this approximate accuracy of response is no guide as to its control mechanism. The issue is also unresolved for the homologous phenomenon in rats (Booth & Davis, 1973). Zbrożyna (1958) established a conditioned inhibitory stimulus in dogs by presenting it at the end of a meal (food rather than its after-effects being the unconditioned stimulus). However, the meal was ended by withdrawal of food, rather than by spontaneous cessation of intake as in the present paradigm.

The experiment was designed to separate subjects who showed differential satiation to the carbohydrate loads on their very first experience of them from the subjects who did not. This was intended to permit an unamibiguous demonstration of response differentiation during acquisition. Indeed, the criteria of selection produced a slight but significant undereating following the concentrated load initially (Table 1, total weight of meal) and so the successful acquisition in this group was an unequivocal crossover effect. Unexpectedly, the subjects who were initially somewhat satiated by the concentrated load did not give the ambiguous result of some increase in this difference and some weak sensory control in extinction. Instead, little or no further differentiation occurred during acquisition, and extinction gave no evidence whatever of acquired sensory control. Several explanations of this result are possible. Individual differences in rapid satiability by concentrated carbohydrate could conceivably correlate with lack of attention to or control by flavour cues. In some studies differences of this sort have been found between lean and obese people, but all subjects in the present experiment were of normal weight. Also, the difference seems rather too large and consistent to relate to such findings. While not excluding the possibility of underlying individual differences, process explanations should be considered. Under conditions where it is relatively easy to discriminate a relationship between sensory and satiating properties of the carbohydrate drinks, this relation could provide cues or control which interfere with attention to or belief in the relation between dessert flavour and differences in satiety. Three-quarters of the subjects in the initially satiated group did indeed guess correctly the relation between sensory and nutrient properties of the drinks, twice the number in the group who acquired control by dessert flavour. Nevertheless there is another and simpler explanation. People in the two groups could differ in gastrointestinal processing of the loads, at least at lunchtime, and a more rapid satiation result in one group. This might prevent the occurrence of an effect such as unusually intense or prolonged satiation after the meal, which is experienced by the other group in which a lunch of unreduced size is consumed on top of a carbohydrate-rich load. That is, the direct effect of the load on intake might eliminate differential consequences of ingesting the loads, leaving nothing to reinforce or to condition the acquisition of responses to correlated sensory differences. As the satiating effects and the intake compensations after loads and meals have yet to be measured precisely in man and are probably imprecise in the short term (Booth *et al.* 1970; Pliner, 1973), it is not possible at present to assess the quantitative adequacy of this explanation.

The group which acquired sensory control showed fewer signs of awareness of potential effects of the loads on appetite and had no recall that the flavoured yoghurts differed in their effects, although it must be noted that the proportion of subjects who reported differences between drinks or yoghurts was insufficient to permit statistical treatment. It remains controversial whether the subject's awareness or learning of the contingency between conditioned and unconditioned stimuli is necessary for all forms of conditioning in man (Dawson, 1973; Mandel & Bridger, 1973). The recall tests were not designed specifically to exclude the possibility that the subjects would have been able under suitable conditions to formulate an impression that, for example, yoghurt of a particular flavour consistently preceded heavy satiation shortly after lunch. The present experiment was designed primarily to determine whether a behavioural phenomenon was shown by human subjects, rather than to elucidate how it might be mediated. The perceptual and subjective aspects of the phenomenon remain to be explored.

Feeding habits are widely acknowledged to be important in appetite. The term 'habit' is, however, notoriously vague and the operative types of behavioural organization should be more precisely specified. One type of feeding control is, we conclude, treating a foodstuff according to past experience of the relationships between its sensory qualities and after-effects of ingestion. The concept of sensory evaluation of foods should be expanded to include satiating power as well as palatability.

This work was done at the Universities of Sussex and Birmingham with support from the British Nutrition Foundation. We thank all the volunteers for their participation.

REFERENCES

BELLOWS, R. T. (1939). Time factors in water drinking in dogs. *Am. J. Physiol.* **125**, 87–97.

BOOTH, D. A. (1972). Conditioned satiety in the rat. *J. comp. physiol. Psychol.* **81**, 457–471.

BOOTH, D. A., CAMPBELL, A. T. & CHASE, A. (1970). Temporal bounds of post-ingestive glucose induced satiety in man. *Nature, Lond.* **228**, 1104–1105.

BOOTH, D. A. & DAVIS, J. D. (1973). Gastrointestinal factors in the acquisition of oral sensory control of satiation. *Physiol. Behav.* **11**, 23–29.

BOOTH, D. A., LEE, M. & MCALEAVEY, C. (In preparation.) Physiological correlates of the satiating effect of concentrated starch solution in man.

BOOTH, D. A., LOVETT, D. & MCSHERRY, G. M. (1972). Postingestive modulation of the sweetness preference gradient in the rat. *J. comp. physiol. Psychol.* **78**, 485–512.

BOOTH, D. A., STOLOFF, R. & NICHOLLS, J. (1974). Dietary flavour acceptance in infant rats established by association with effects of nutrient composition. *Physiol. Psychol.* **2**, 313–319.

CABANAC, M., PRUVOST, M. & FANTINO, M. (1973). Alliesthesie negative pour des stimulus sucrés après diverses ingestions de glucose. *Physiol. Behav.* **11**, 345–348.

CHURCH, C. F. & CHURCH, H. N. (1970). *Food Values of Portions Commonly Used: Bowes and Church*, 11th ed. Philadelphia: Lippincott.

DAVIS, J. D. & CAMPBELL, C. S. (1973). Peripheral control of meal size in the rat: Effect of sham feeding on meal size and drinking rate. *J. comp. physiol. Psychol.* **83**, 379–387.

DAWSON, M. E. (1973). Can classical conditioning occur without contingency learning? A review and evaluation of the evidence. *Psychophysiology* **10**, 82–86.

HOLMAN, E. (1973). Temporal properties of gustatory spontaneous alternation in rats. *J. comp. physiol. Psychol.* **85**, 536–539.

HOLMAN, G. L. (1969). Intragastric reinforcement effect. *J. comp. physiol. Psychol.* **69**, 432–441.

HSIAO, S. & TUNTLAND, P. (1971). Short-term satiety signals generated by saccharin and glucose solutions. *Physiol. Behav.* **7**, 287–289.

JORDAN, H. A., STELLAR, E. & DUGGAN, S. Z. (1968). Voluntary intragastric feeding in man. *Commun. behav. Biol.* **1**, 65–69.

LE MAGNEN, J. (1955). La satiété induite par les stimuli sucrés chez le rat blanc. *C. r. Séanc. Soc. Biol.* **149**, 1339–1342.

LE MAGNEN, J. (1956*a*). Hyperphagie provoquée chez le rat blanc par alteration du mécanisme de satiété peripherique. *C. r. séanc. Soc. Biol.* **150**, 32–35.

LE MAGNEN, J. (1956*b*). Rôle de l'odeur ajoutée au régime dans la régulation quantitative à court terme de la prise alimentaire chez le rat blanc. *C. r. Séanc. Soc. Biol.* **150**, 136–139.

LE MAGNEN, J. (1956*c*). Effets sur la prise alimentaire du rat blanc des administrations postprandiales d'insuline et le mécanisme des appétits caloriques. *J. Physiol., Paris* **48**, 789–802.

LE MAGNEN, J. (1959). Effets des administrations postprandiales de glucose sur l'établissement des appétits. *C. r. Séanc. Soc. Biol.* **153**, 212–215.

LE MAGNEN, J. (1961). Effets d'une pluralité de stimuli alimentaires sur le determinisme quantitatif de l'ingestion chez le rat blanc. *Arch. Sci. physiol.* **14**, 411–419.

LE MAGNEN, J. (1969). Peripheral and systemic actions of food in the caloric regulation of intake. *Ann. N.Y. Acad. Sci.* **157**, 1126–1156.

MANDEL, I. J. & BRIDGER, W. H. (1973). Is there classical conditioning without cognitive expectancy? *Psychophysiology* **10**, 87–90.

McCance, R. A. & Widdowson, E. M. (1969). *The Composition of Foods*. London: HMSO.

Miller, N. E. (1957). Experiments on motivation. *Science, N.Y.* **126**, 1271–1278.

Pliner, P. L. (1973). Effect of liquid and solid preloads on eating behaviour of obese and normal persons. *Physiol. Behav.* **11**, 285–290.

Richter, C. P. (1942–3). Total self-regulating functions in animals and human beings. *Harvey Lectures* **38**, 63–103.

Rozin, P. & Kalat, J. W. (1971). Specific hungers and poison avoidance: adaptive specializations of learning. *Psychol. Rev.* **78**, 459–486.

Shama, F. & Sherman, P. (1973). Identification of stimuli controlling the sensory evaluation of viscosity. II. Oral methods. *J. Texture Studies* **4**, 111–118.

Simson, P. C. & Booth, D. A. (1973). Olfactory conditioning by association with histidine-free or balanced amino acid loads. *Q. Jl exp. Psychol.* **25**, 354–359.

Teitelbaum, P. & Epstein, A. N. (1963). The role of taste and smell in the regulation of food and water intake. In Y. Zotterman (ed.), *Olfaction and Taste*. New York: Macmillan.

Wood, F. W. (1968). Psychophysical studies on the consistency of liquid food. *Soc. Chem. Ind. Monogr.* no. 27, 40–49.

Zbrożyna, A. W. (1958). On the conditioned reflex of the cessation of the act of eating. II. Differentiation of the conditioned cessation reflex. *Acta Biol. exp.* **18**, 163–166.

(*Manuscript received* 11 *November* 1974)

Br. J. Psychol. (1976), 67, 2, *pp.* 149–160
Printed in Great Britain

SYNTACTIC CONSTRUCTIONS AS RETRIEVAL PLANS

BY GERARD KEMPEN
Experimental Psychology Unit, University of Nijmegen, Nijmegen, The Netherlands

Four probe latency experiments show that the 'constituent boundary effect' (transitions between constituents are more difficult than within constituents) is a retrieval and not a storage phenomenon. The experimental logic used is called paraphrastic reproduction: after verbatim memorization of some sentences, subjects were instructed to reproduce them both in their original wording and in the form of sentences that, whilst preserving the original meaning, embodied different syntactic constructions.

Syntactic constructions are defined as pairs which consist of a pattern of conceptual information and a syntactic scheme, i.e. a sequence of syntactic word categories and function words. For example, the sequence noun + finite intransitive main verb ('John runs') expresses a conceptual actor–action relationship. It is proposed that for each overlearned and simple syntactic construction there exists a retrieval plan which does the following. It searches through the long-term memory information that has been designated as the conceptual content of the utterance(s) to be produced, looking for a token of its conceptual pattern. The retrieved information is then cast into the format of its syntactic scheme. The organization of such plans is held responsible for the constituent boundary effect.

It is generally assumed that long-term memory representations of sentences or text are conceptual in nature unless the text has been memorized by subjects who anticipated a verbatim recall test. Then, in addition, the literal wording of the sentences is stored. Indeed, several phenomena of sentence recall are easy to explain by appealing to some sort of verbatim storage. One of these is the 'constituent boundary effect': word-to-word transitions *between* syntactic constituents (of a certain level) are more difficult than such transitions *within* constituents (of the same level). A convenient after-learning measure of transitional difficulty is probe reaction time; subjects are presented with one word (probe) of a memorized sentence and respond as fast as possible by saying the next word of the same sentence (Suci *et al.* 1967; Kennedy & Wilkes, 1968).

The constituent boundary effect – this term is intended to be descriptive, not interpretative – has invariably been explained in terms of storage, i.e. by postulating a type of memory storage that favours the combined access to both members of certain word pairs (e.g. two words belonging to the same syntactic constituent) but not of other pairs (e.g. two words from different constituents). A well-known version of this principle includes the notion of hierarchical chunking or grouping (Johnson, 1965, 1970; Wilkes & Kennedy, 1969; see also next section). However, the constituent boundary effect and, generalizing to non-sentence sequences of words or letters, grouping effects do not preclude alternative interpretations in terms of retrieval processes. On this type of account, transitional difficulties are attributed to the organization of a retrieval plan which at certain stages of its execution engages in more elaborate processing activity for locating and outputting the next word of a learned sequence, than at other stages. Johnson (1970), writing in the context of experiments on learning grouped letter strings, remarks that these alternative explanations are empirically indistinguishable because the usual experimental logic for assessing the relative importance of storage and retrieval does not apply.

This logic, if applied to profiles of transition difficulties, would consist of having the subjects retrieve the memorized sequence by means of a retrieval plan that presumably imposes a different profile on to the sequence than the profile to be expected from hypothetical groups or chunks in the memory store. The extent to which the empirical profile obtained under such a retrieval plan instruction conforms to the profile expected from the organization of this plan, determines the plausibility of the retrieval interpretation. In the case of non-sentence strings this method is obviously inapplicable due to the absence of alternative retrieval plans.

However, this limitation does not necessarily exist for sentences. The present paper presents evidence which, on one hand, disconfirms storage interpretations of the constituent boundary effect and, on the other hand, supports a specific retrieval interpretation. The experiments to be reported indicate that, closely related to syntactic constructions, there exist retrieval plans which are capable of searching through conceptual representations of sentences and casting the retrieved information into a specific output format for reproduction purposes. The existence of these syntactic retrieval plans is relevant in two respects. First, they embody a principle of memory search and retrieval that has not yet been described. Secondly, they allow a realistic view on the memory processes during verbatim learning and recall of sentences and text. In current explanations of the constituent boundary effect, constituents are treated on a par with item groups in non-sentence strings. Apart from leaving open the question why, in the case of sentences, subjects often adopt groupings which correspond to syntactic constituents, such explanations fail to recognize the possibility that subjects derive their responses from a memory representation of the sentences' meanings. The to-be-memorized sentences typically are easy to understand and it seems realistic to allow for storage of conceptual representations and for their use during verbatim recall tasks.

SYNTACTIC RETRIEVAL PLANS

Suppose the subjects have committed to memory a few sentences of the type (1) *Two Finns wrote texts* or (2) *Three Greeks read books*. A storage theory of the constituent boundary effect might hypothesize that the long-term memory representation of (1) is the following hierarchy of chunks or registers:

(3) (S (NP (Q *two*) (N *Finns*)) (VP (V *wrote*) (N *texts*)))

The topmost level of the hierarchy is occupied by the register S which contains the symbol sequence 'NP VP'. NP itself is the name of a register with the sequence 'Q N' as its content. At the lowest level there are registers (e.g. Q and N) which always contain one single *terminal symbol* such as *two* and *Finns*.

There are several possibilities for devising a procedure which will reconstruct the memorized sentence from such hierarchies. One of those, which is very similar to the mechanism described by Johnson (1970), will be outlined here. The operations to be performed are controlled by the list of symbols called the pushdown stack, and by the symbol in an auxiliary location called Aux. The sole accessible symbol of the stack is its topmost element, T; the only permissible stack alterations are placing a new symbol on top of the stack and removing T from the stack. Of course, both types of alterations result in a new T. The procedure has been labelled Reprod and

Table 1. *Essentials of a possible procedure for reconstructing a learned sentence or word string from a hierarchy of labelled chunks or registers*

(For explanation see text.)

Reprod

```
(1) If T is a terminal symbol then pronounce T
    else copy T into Aux
(2) Remove T from stack
(3) Fillstack
```

Fillstack

```
(1) If reg(Aux) is not empty then
        (a) copy the last symbol of reg(Aux) on
            to stack
        (b) remove this symbol from reg(Aux)
    else Reprod
(2) Fillstack
```

it includes the subroutine Fillstack; an occurrence of any of these labels in a program statement means that the associated (sub)procedure is called. Table 1 contains the essential part of Reprod.

Reprod starts with the label of the highest-order register (that is, S, in case of hierarchy (3)) on the stack. The assumption has been made that once a symbol has been copied from a register onto the stack it is removed from the register (statement 1*b* of Fillstack). This obviously irrealistic but inessential and simplifying feature is easily amendable. Since T is not a terminal symbol it is copied into Aux and control is transferred to Fillstack. Fillstack puts the symbols of the register which is indicated by Aux (reg(Aux)) onto the Stack in such a way that the first register symbol is T at the moment control is returned to Reprod. With NP figuring as T, now, Reprod calls Fillstack again after having written 'NP' into Aux. Fillstack moves the NP content to the stack, etc. Only after *two* has been placed on top of the stack Reprod will utter the first response word. It is easy to see that if two consecutive responses belong to the same constituent, the amount of symbol manipulation carried out is smaller than if they belong to different constituents: the basis of the constituent boundary effect.

Sentences (1) and (2) are instances of the syntactic construction Quantifier + Noun + Finite main verb + Noun, with the first noun serving as head of the subject phrase, the second noun as object, and the sequence Main verb + Noun as predicate phrase. Now consider strings such as (4) *Two Finns*, (5) *Finns wrote* and (6) *wrote texts*. Each of them, too, is an instance of a grammatical construction and expresses part of the meaning of (1). For convenience, and somewhat stretching the meaning of the term, let us call (4)–(6) *partial paraphrases* of (1). A syntactic construction may be defined as a sequence of syntactic categories and function words which expresses one or more conceptual case of modifier relationships. For short, a syntactic construction is a pair consisting of a syntactic scheme and a conceptual pattern. For example, the Noun + Finite intransitive main verb construction (*Finns wrote*) expresses agentive case relationships between an unmodified actor and an unmodified action. From this definition it is not difficult to see how syntactic constructions could work, in principle, as devices for retrieving information from memory representations and map it into syntactic schemes as output formats. One needs to assume that, for each syntactic construction that is heavily overlearned and relatively simple, the subjects have available a procedure which searches through a designated area of long-term memory – the area which contains the conceptual content of the utterance(s) to be produced. There it looks for a token of its conceptual pattern and transforms the collected information into its output format (consulting, whenever necessary, the internal lexicon).

How can these syntactic retrieval plans account for the constituent boundary effect? Suppose a retrieval plan has located a piece of information which fits into its conceptual pattern. At this moment, the retrieval plan initiates its second task: assembling a word sequence according to the specifications of its syntactic scheme. We will assume that the procedure which performs this task consists of a hierarchy of subprocedures. Take hierarchy (3), for instance. Now, the non-terminal symbols are to be interpreted as labels of subprocedures which assemble a specific type of constituent, and no longer as names of already filled registers or chunks. The terminal symbols indicate the results obtained by the subprocedures. The program of Table 1 governs the order in which the various subprocedures are called and, again, will produce a constituent boundary effect.

Returning to sentence (5), the procedure which fabricates constructions of this type might be represented by (S′(N—) (*V*—)), with dashes referring to open places of the output format. Apart from the identity of the non-terminal symbols, this hierarchy is identical to the NP and VP parts of (3).

The probe latency technique entails two deviations from the usual way of sentence reproduction: the reproduction sequence does not necessarily start with the first word of the sentence, and it consists of one word only. Without working out the details it will be assumed that presentation of a probe word sets the current retrieval plan to the state which during normal sentence reproduction would have been present at the moment of pronouncing that word. Furthermore, it is to be supposed that the subjects, in response to the task instructions given to them, have set up a control mechanism which forces the retrieval plan to halt after the first response word has been uttered.

PARAPHRASTIC REPRODUCTION: EXPERIMENTAL PREDICTIONS

In order to apply the previously mentioned experimental strategy for solving storage vs. retrieval problems (see introductory section) to sentence memory one can exploit the availability of partial and full paraphrases of the memorized sentences. If the retrieval interpretation is right and the subjects derive responses from conceptual representations of the learned sentences, then, during appropriately devised recall tests, they must be able to use the retrieval plans which correspond to (partial) paraphrases of the learned sentences. Interpretation of profiles of transition difficulties yielded by this *paraphrastic reproduction* method proceeds in accordance with the general storage vs. retrieval logic. In sentences (1) and (2), the middle transition requires a greater amount of retrieval plan activity than both neighbouring transitions and is expected, other things being equal, to take more time. Now take an experiment where the subjects, after having learned four sentences of type (1) or (2), are told that during the subsequent probe reaction time trials only probe and response words from the four partial paraphrases such as (5) will occur. This represents a situation where the subjects can profitably apply the retrieval routine associated with a familiar construction which is simpler than the construction of the learned sentences. This consideration, however, generalizes to the transitions in partial paraphrases (4) and (6) if their difficulties would be measured in the analogous way.

In this new type of task a general decrease in average latency is to be expected because of the smaller number of possible probe and response items. Since none of the constructions exemplified by (4), (5) and (6) contain compound constituents (of word length two or more), the retrieval hypothesis would expect the mean latency for the middle transition to decrease more sharply than the other transitions when comparing the usual way of measuring probe latencies to this new way. (One cannot make the stronger prediction that the three decrease to the same level of difficulty since the *ceteris paribus* condition is not necessarily true: the words flanking the three transitions are different, and the three subroutines might take different amounts of time.)

As for terminology, the usual method of probe reaction time measurement, where probes are taken at random from the various transitions, will be called 'sentencewise'. The alternative paradigm with many successive probes belonging to the same transition, will be called 'pairwise'.

The interaction predicted by the retrieval hypothesis (the latency decrease as a consequence of shifting from the sentencewise to the pairwise paradigm is greater for between-constituent transitions than for within-constituent transitions) must not appear, however, in sentences in which the between-constituent transition is not surrounded by a word pair making up a syntactic construction. In the Dutch language such a condition can easily be arranged without altering the words from which sentences like (1) and (2) have been composed.

Word order within the Dutch predicate constituent is different in main and subordinate clauses. In main clauses the finite verb precedes other predicate parts such as object, adverb and prepositional phrase; in subordinate clauses this order is reversed. For instance, if the Dutch equivalent of (1) is embedded as a subordinate clause within a main clause of the type *He says that...*, the result would be – in a word-to-word translation – (6) *He says that two Finns texts wrote*. Here, the between-constituent transition is no longer flanked by a partial paraphrase because (7) *Finns-texts* is not a syntactic construction. Hence, the above interaction between paradigm and transition should not appear.

In this set of predictions the interaction for type (1) sentences is critical with regard to the storage vs. retrieval issue. The storage assumption predicts that the constituent boundary effect will remain constant through variations in the measurement paradigm. In the following experiments these predictions were put to test, together with a control that allowed elimination of a trivial short-term memory interpretation of the outcomes.

EXPERIMENT I

The first experiment is basically a replication, with two modifications, of the traditional constituent boundary effect as measured with the sentencewise probe latency technique. Many of the procedural details also apply to Expts. II–IV.

The first modification is the simultaneous measurement of probe latencies for *more than one* sentence. In earlier studies, sentences were probed one-by-one and subjects had the opportunity to concentrate on one sentence only. The pairwise paradigm, however, presupposes that several sentences are simultaneously kept in memory. In order to equalize the sentencewise and the pairwise paradigms in all non-critical aspects it was decided that in the sentencewise condition, too, a set of four sentences should be memorized and probed concurrently. A pilot experiment with the sentencewise paradigm pointed out that if the subjects have advance knowledge of the sentence which is about to be probed, the pattern of results is not affected, except for a general decrease in mean latency.

The second modification consists of the introduction of backward responses, i.e. of a condition where the subjects have to name the predecessor, not the successor, of the probe. The purpose of this modification is not merely to add a few data points: it yields the opportunity to test the hypothesis that the constituent boundary effect should be more pronounced in the backward than in the forward reactions. This hypothesis follows from the conception of the retrieval system described above together with the reasonable assumption that any *ad hoc* change in a previously acquired plan slows down the reactions that necessitated that change.

In order to change the program of Table 1 in such a way that it will reproduce a string in reversed order, the only amendment needed is to replace the word 'last' in statement 1*a* of Fillstack by 'first'. In case of a within-constituent transition this statement is executed less often than in case of a between-constituent transition. Taking the hierarchy of (3) for instance, statement 1*a* is encountered only once between responses *two* and *Finns* but three times in between *Finns* and *wrote*. This means that a between-constituent transition, which is already difficult by itself, will suffer more from the requirement of direction reversal.

Method

Subjects were 20 students at the University of Nijmegen who took part in individual sessions. Each subject learned a different set of four Dutch sentences which had been composed from the following lexical material. The first word of all four sentences was *Die* (Those). The second position was occupied by the numerals *twee, drie, vier* and *vijf* (two, three, four, five). The third words were *Belgen, Finnen, Grieken* and *Ieren* (Belgians, Finns, Greeks, Irish). In fourth position occurred the past tense verbs *lazen, leerden, schreven* and *typten* (read, learned, wrote, typed). They belong to the category of middle verbs which can serve both transitive and intransitive function. The final words were selected from four different categories, one word from each category. (The reason for this complication has to do with a theoretical issue that will be touched on in the footnote only.) The categories were Objects (*boeken. teksten. woorden, zinnen*; books, texts, words, sentences), Adjectival Adverbs (*keurig, matig, moeizaam, prettig*; nicely, moderately, laboriously, pleasantly), Time Adverbs (*'s avonds, 's morgens, 's winters, 's zomers*; in the evening, in the morning, in the winter, in the summer), and Frequency Adverbs (*altijd, dikwijls, nimmer, zelden*; always, often, never, seldom). Each word sequence that arises by arbitrary selection of one word from each of the five position groups constitutes a grammatical sentence.

The 20 four-sentence sets were systematically made up in a way that guaranteed maximal between-set variation in the concrete to-be-learned sentences. Learning material of one subject could be, for example, the set ***Die drie Finnen typten moeizaam, Die twee Grieken leerden boeken, Die vijf Ieren schreven 's avonds, Die vier Belgen lazen dikwijls.***

The sentences were learned from a memory drum that exposed a new line of the learning material every four seconds, by means of a study–test method. The recall stimuli were the first two words of the sentences, e.g. *Die twee, Die drie*. Memorization went on till the criterion of word perfect recall and, in addition, there was 100 per cent overlearning.

Probe reaction times were measured by a voice-key that started when the probe word appeared in the window of a memory drum and stopped at the onset of subject's pronouncing the response word. The interval between consecutive probes was 6 sec. Reaction times of 3000 msec or longer

Table 2. *Average latencies (arithmetic means) for the two main transitions in Expts. I–IV*

	Transition and direction			
	Between-constituents		Within-predicate	
Experiment	Forward	Backward	Forward	Backward
I	1725	2061	1561	1591
II	1339	1454	1378	1382
III	1446	1469	1335	1376
IV	1368	1654	1258	1301

were discarded as errors. The instruction required the subject to name, as quickly as possible upon the presentation of a word from one of the memorized sentences, avoiding errors as much as possible, the word that immediately followed the probe word in the sentence concerned (or preceded it, in the case of backward reactions). If the probe was the last word of a sentence subjects had to reply with 'punt' ('stop'); in the backward condition, the pronoun *Die* (Those) served as the analogue of 'punt'. Each subject took part in both forward and backward conditions, the order of these conditions being systematically varied to the effect of controlling for practice influences. The list of probes for a subject contained four occurrences of each of the 16 possible probes in random order, and there were separate lists for the backward and the forward conditions. Both started with four arbitrary training probes. Each subject produced a maximum of $4 \times 16 \times 2 = 128$ latency scores. During the learning stage of the experimental sessions subjects were unaware of the fact that a probe latency task would follow.

Results

Due to the following sources of erroneous responses, not all intended observations could be obtained: incorrect answers; no response within 3000 msec; answers given without sufficient voice volume to stop the timer; and premature stopping of the timer as a result of irrelevant noises made by subject, apparatus or environment. In this paper the data analysis will be confined to the middle and the right-hand transitions, that is, to the between-constituent transition and the transition within the predicate. The left-hand transition (such as *Twee Grieken*) will not be considered because the quantifiers served as recall stimulus during learning. The transitions with respectively 'stop' and 'Those' as responses are even more atypical, for obvious reasons.

Table 2 contains the mean reaction times for the two main transitions and for both forward and backward reactions. The percentage of errors here was 14; they were excluded from further analysis. The constituent boundary effect amounted to 164 msec in the forward direction; it was much stronger in the backward direction: 470 msec. The upper pair of dotted lines in Fig. 1 represents the same data. An analysis of variance confirmed that the Transition by Direction interaction which is at stake here was highly significant ($F = 74{\cdot}5$; d.f. $= 1, 19$; $P < 0{\cdot}0001$). Application of the t test for *a priori* comparisons to the 164 msec difference between the two forward means yielded a value of $t = 4{\cdot}65$ (d.f. $= 19$; $P < 0{\cdot}0005$).

The conclusion is that the procedure of this experiment favours the appearance of the constituent boundary effect, especially in the backward condition. The next experiment concentrates on the pairwise paradigm.

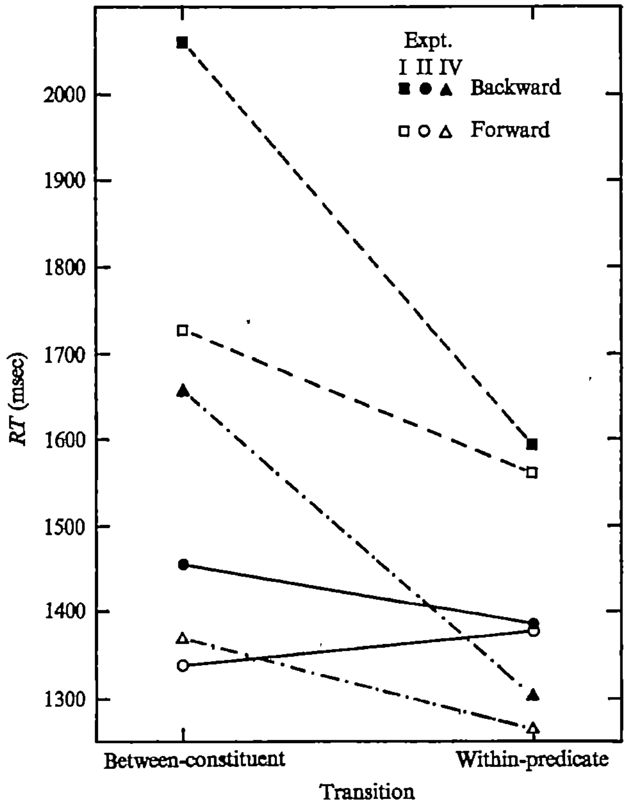


Fig. 1. Mean latencies for Expts. I, II and IV. Expt. I: sentence wiseparadigm. Expt. II: pairwise paradigm with partial paraphrase at between-constituent transition. Expt. IV: pairwise paradigm but no partial paraphrase at between-constituent transition.

EXPERIMENT II

Measurement of probe latencies according to the pairwise paradigm in which consecutive probes come from one type of transition creates a situation where the subjects have advance knowledge of the transition being tested, and affords them the opportunity to mobilize a syntactic retrieval plan specifically tuned to that transition, at least if the word pairs surrounding that transition are instances of a syntactic construction. The latter condition is fulfilled by the sentence material of Expt. I because all three transitions are flanked by tokens of syntactic constructions, e.g. *two–Finns*, *Finns–wrote* and *wrote–texts*.

The pairwise paradigm may be expected to entail a general increase in the speed of reactions since it curtails the number of possible probes and responses substantially as compared with the sentencewise paradigm. But if the subjects bring into play specific retrieval routines for the construction exemplified by *Finns wrote*, then this transition no longer suffers from the disadvantage of being a transition between compound constituents. Hence, the reaction times for the between-constituent transition are predicted to improve more than those within constituents.

As in Expt. I, both backward and forward measures were taken, though the two types were intermingled. That is, not only *Finns* but also *wrote* served as a probe within the same series of trials, and the subjects were instructed to respond with the other member of the pair.

Method

Apart from the paradigm of probe latency measurement the procedure of the present experiment was identical to the previous one. As a matter of fact, the two were run concurrently and 40 subjects were alternately assigned to either of them. The list of probes was divided into four parts, one for each transition measured, namely the between-constituent transition, the transition within the predicate phrase and two transitions within pairs of non-adjacent words, e.g. *Finns–texts*. For each transition there were eight possible probes: the left-hand and the right-hand members of the four surrounding word pairs; they required forward and backward responding, respectively. Each probe word occurred four times on the list. The total number of trials was 4 (transitions) × 8 (probes per transition) × 4 (repetitions of probe words) = 128, as in Expt. I.

The instruction for each transition was communicated to the subject in terms of word pairs occupying certain serial positions in the memorized sentences. For example, 'Each of the following 32 probes is either the third or the fourth word of a sentence; if you see the third word, respond with the fourth word of the same sentence, and vice versa.' In order to control for practice effects the presentation order of the transitions was systematically varied over subjects. They heard about the probe latency measurement only after the learning stage.

Results

The following data do not include 5 per cent of erroneous responses. The main latencies as shown in Table 2 and in the continuous lines of Fig. 1 make amply clear that the between-constituent transition took more advantage of the shift to the pairwise paradigm than the within-constituent transition. The constituent boundary effect has almost disappeared: only in backward direction a weak effect of 72 msec has remained ($t = 2{\cdot}00$; d.f. $= 19$; $P < 0{\cdot}05$; t test for *a priori* comparisons), sharply contrasting with the 470 msec effect of Expt. I. The slight 'negative' constituent boundary effect for the forward reactions is not statistically reliable.

DISCUSSION

Although these results accord with the predictions, they are open to a trivial interpretation. One could hypothesize that during the empty intervals between a response and the next probe the subjects engaged in rehearsing the four word pairs from which the probes were selected and thus were responding on the basis of information kept in short-term or working memory and not by interrogating the long-term memory representations of the learned sentences. If this were true, one would no longer expect the between-constituent transitions to take more time than the other ones. This calls for a control experiment which prevents the subjects from implicit rehearsal during empty intervals.

EXPERIMENT III

In order to ascertain that, at the moment of presentation of a probe, the subject's working memory does not contain the four word pairs which are being tested during a given series of probes, a filler task was introduced that consisted of reading aloud, at high speed, a list of consonant letters. The subjects started this activity as soon as they had produced a response or, in case of no response, upon a signal that the maximum allowed latency of 3000 msec had elapsed. They went on reading until an acoustic signal warned them to look back to the position of the memory drum window where the next probe would appear. The letters were visible in the said memory drum window, to the right of the probe position. The warning signal sounded 700 msec before the next probe was due. There were eight new subjects from the same general population as before. Further procedural details were identical to those in Expt. II, except that all subjects learned the same set of sentences.

The pattern of results (cf. Table 2; 9 per cent errors were discarded) was in good agreement with that of the previous experiment. The means of the backward reactions were even nearly identical to their counterparts in Expt. II. As for the forward reactions, the slight negative constituent boundary effect in Expt. II had changed to a positive effect of 111 msec which, however, was not significant ($t = 1{\cdot}46$; d.f. $= 7$; t test as above). This might be interpreted as discomforting to the retrieval plan hypothesis because the effect was not very much stronger in the sentencewise paradigm (164 msec, Expt. I). However, since the profile of the backward latencies has been left untouched by the short-term memory control and since, after all, the forward constituent boundary effect is smaller, though not significantly so, than in Expt. I, it is concluded that the short-term memory explanation is not valid. The reader will recall the theoretical and empirical arguments (Expt. I) that backward reactions are liable to stronger constituent boundary effects than forward ones. This justifies more weight to be attached to backward than to forward reactions. The question why the filler task does affect the latter but not the former reactions remains open. (The results of Expt. IV present further arguments against a straightforward short-term memory interpretation; cf. footnote.)

EXPERIMENT IV

It remains to be demonstrated that in the pairwise paradigm the pattern of outcomes typical for the constituent boundary effect will reappear when the word pairs at both sides of the between-constituent transition do not make up a partial paraphrase of the memorized sentences (cf. the section entitled Paraphrastic Reproduction). One possible way of generating sentences fulfilling this requirement is simply to convert the sentences of the previous experiments into subordinate clauses, for instance, by prefacing them with the conjunction *because*. This yields, in a word-to-word Dutch to English translation, clauses such as *Because those two Finns texts wrote...*, with the pair *Finns–texts* at the boundary between constituents.

One further modification of the sentences was needed. In Expts. I and II, the words following the subject nouns belonged to one syntactic category (verbs), but the words in final position were members of varying parts of speech (nouns, adverbs). Reversing the predicate word order would move the latter to the position following the subject nouns. In order to enhance the comparability of the present experiment with Expt. II it was decided to use only one type of predicate construction, namely Object+Main Verb, so that all words following the main constituent boundary belonged to one syntactic category (see Expt. I for a listing of the object nouns).

Further procedural details were exactly the same as in Expt. III, although there was no filler task. Eight new subjects had been recruited from the same population.

The main results (16 per cent errors were discarded) are shown in Table 1 and by the lower pairs of dotted lines in Fig. 1. It is obvious that the general pattern of the constituent boundary effect has been restored. In the backward direction the effect was again very strong. In the forward direction it is significant ($t = 2{\cdot}23$; d.f. $= 7$; $P < 0{\cdot}05$, t test as above), although the between-constituent transition took only 110 msec more than the transition within the predicate phrase. It is also instructive to look at the latencies obtained for the non-adjacent word pairs. In this study, as in Expts. II and III, reactions were measured over the pairs formed by subject nouns and final words. In the previous studies the non-adjacent pairs always produced longer latencies than the adjacent pairs, the difference typically being 150 msec or more (averaged over forward and backward responses). Such an effect would be predicted by the syntactic retrieval hypothesis because non-adjacent pairs like *Finns–texts* or *Greeks–always* do not comprise syntactic constructions, let alone partial paraphrases. In the present experiment, however, the non-adjacent pairs do constitute partial paraphrases (*Finns wrote*, etc.) and one would expect considerable facilitation as a result of this. Indeed, the non-adjacent pairs now take 1350 msec on the average, which is 160 msec below the average of the adjacent between-constituent transition. The conclusion must be that the latency profiles behave in good agreement with the predictions derived from the notion of syntactic constructions as retrieval plans.*

GENERAL DISCUSSION

The first implication of this series of paraphrastic reproduction experiments is that the frequently observed constituent boundary effect in verbatim sentence recall comes into existence at retrieval time. Hierarchically organized retrieval plans provide a sufficient explanation and it need not be assumed that the sentence representations which are accessed by these mechanisms are divided into parts corresponding to syntactic constituents. Plausible candidates for this type of sentence representation are conceptual representations in the form of predicates over arguments or cases, or of imagery codes. This contrasts with the dominant view that verbatim recall largely reflects properties of the phonetic or surface structure representations which the subjects supposedly store while memorizing. The present results indicate that rote learning of sentences gives rise to easily accessible and retrievable memory representations of their conceptual contents, perhaps in addition to phonetic representations. Other proponents of retrieval interpretations of sentence recall phenomena are Loosen (1972) and James *et al.* (1973). The latter group of investigators take this stand in the context of another so-called syntactic effect in sentence memory: the 'tendency towards the kernel' as found in the recall of active, passive, negative, etc., sentences (e.g. Mehler, 1963).

---

* Two remarks on the interpretation of Expt. IV. First, it should be noticed that the results of Expt. IV refute the straightforward STM explanation outlined in Expt. III, but they remain compatible with a more limited STM interpretation, namely that the subjects were rehearsing word pairs that constitute syntactic constructions (and no other word pairs). This version is also contradicted by the results of Expt. III, but, if it were true, it would have to add some account of why syntactic constructions are so easy to rehearse!

Second, assuming that sentences receive memory codes in the form of conceptual networks (e.g. *Finns*–actor–*write*–object–*texts*), one might argue that the distance between *Finns* and *wrote* is smaller than between *Finns* and *texts*. This might provide an alternative explanation of the different reaction time profiles of Expts. II and IV. However, this line of reasoning runs counter to another aspect of the data that has not been mentioned yet. If distance over nodes of a conceptual network were an important determinant of reaction times, one would predict substantial effects of the various grammatical relations in the predicate phrase. Remember that, apart from (*a*) object nouns, there were (*b*) adjectival adverbs (verb modifiers), (*c*) time adverbs, and (*d*) frequency adverbs. In conceptual network representations, (*a*) and (*b*) would be more closely connected to the verb than (*c*) and (*d*), that serve as sentence modifiers. But in none of the three experiments did type of grammatical relation produce effects that were statistically reliable. Moreover, the effects changed direction over experiments.

Second and more important, the results point at the existence of long-term memory retrieval routines closely allied with (highly overlearned and simple) syntactic constructions; i.e. routines which scan sentence representations, read out only a specific pattern of conceptual information and cast this into a specific syntactic output format. This contrasts with the view that sentence production proceeds by first deciding on the content of the utterance, e.g. by somehow retrieving conceptual information from memory and, subsequently, mapping this information into a syntactic scheme.

It is suggested that syntactic factors play a role already during the first stage of sentence production, namely as one of the mechanisms for guiding the memory retrieval of conceptual content.

The author is indebted to W. J. M. Levelt and H. Kolk for their valuable criticisms and to Marieke van den Heuvel, F. van Well and R. Schreuder for experimental assistance.

REFERENCES

FILLENBAUM, S. (1973). *Syntactic Factors in Memory?* The Hague: Mouton.

JAMES, C. T., THOMPSON, J. G. & BALDWIN, J. M. (1973). The reconstructive process in sentence memory. *J. verb. Learn. verb. Behav.* **12**, 51–63.

JOHNSON, N. F. (1965). The psychological reality of phrase-structure rules. *J. verb. Learn. verb. Behav.* **4**, 469–475.

JOHNSON, N. F. (1970). The role of chunking and organization in the process of recall. In G. H. Bower (ed.), *The Psychology of Learning and Motivation*, vol. 4. New York: Academic Press.

KENNEDY, R. A. & WILKES, A. L. (1968). Response-times at different positions within a sentence. *Q. Jl exp. Psychol.* **20**, 390–394.

LOOSEN, F. (1972). Cognitieve organisatie van zinnen in het geheugen. (Dissertation, Louvain.)

MEHLER, J. (1963). Some effects of grammatical transformation on the recall of English sentences. *J. verb. Learn. verb. Behav.* **2**, 346–351.

SUCI, G. J., AMMON, P. & GAMLIN, P. (1967). The validity of the probe latency technique for assessing structure in language. *Lang. Speech* **10**, 69–80.

WILKES, A. L. & KENNEDY, R. A. (1969). Relationships between pausing and retrieval latency in sentences of varying grammatical form. *J. exp. Psychol.* **79**, 241–245.

(*Manuscript received* 28 *June* 1974; *revised manuscript received* 2 *April* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 161–163
Printed in Great Britain

RECENCY IN FREE RECALL: STM PLUS LTM OR STM ALONE?

By P. E. MORRIS and J. RIDGWAY
Psychology Department, University of Lancaster

Although alternative views exist (Murdock, 1974), it is generally assumed that the recency component of the serial position curve for free recall is largely the result of recall from short-term memory (STM), while the middle and early items from the list are recalled solely from long-term memory (LTM). Waugh & Norman (1965) attempted to estimate the relative contributions of STM and LTM to the recency effect. They assumed that the last few items can be recalled from either STM or LTM, and that the probability of recall from each is independent of the probability of recall from the other.

However, Broadbent (1971) has argued that 'the probability of recall of these items near the end of the list is constant whatever the length of the list...and this is true for the last, say five or six items of the list' (p. 355). This statement was echoed by Watkins (1974), who stated that 'the variables affecting the recall of prerecency items leave the recency effect invariant, or affect it only slightly' (p. 695). Both authors quote the data of Murdock (1962), which consist of serial position curves for lists of different lengths and different presentation rates. Broadbent (1975) pointed to the overlap in the recency curves for the two 20-item lists, which represent presentation rates of 1 and 2 sec per item. Since the slower presentation rate shows a bigger LTM component, the recency curve for this list should be above that for the faster presentation rate condition. The overlap is taken as evidence against Waugh & Norman's model, and a general conclusion that LTM does not affect recall from the last few items is drawn.

Is the recency effect attributable to STM, or to both STM and LTM? Inspection by eye of partial data seems an inappropriate test of an issue of considerable theoretical interest.

Two predictions from the views of Broadbent and Waugh & Norman were made, and were tested against Murdock's data (reported in Norman, 1970). If Broadbent is correct, then the probability of recall of the items in the recency part of the serial position curve should be independent of the level of recall from the earlier part of the list, and should depend solely on the position of the item relative to the end of the list. For example, the probability of recall of the last but one item should be the same for all lengths of list or rates of exposure. Waugh & Norman predict that the better the recall in the early items of the list, the better the recall in the recency component of the curve.

They assume that R_i, the observed probability of recall of the ith item from the end of the list, is given by

$$R_i = S_i + L_i - S_i L_i,$$

where S_i and L_i represent the probability of recall from STM and LTM respectively.

L_i is estimated by averaging probabilities from the flat, middle part of the curve. Clearly, the model predicts that high values of L_i produce high values of R_i; further, serial position curves for different levels of L_i come together as S_i increases, without intersecting.

To test between these predictions, the probability of recall of items in the recency curve from the six conditions reported by Murdock were examined. The final items were excluded because of the ceiling effect, and because this item is least sensitive to differences between the two models. The probability of recall of the next five items was examined using a Friedman two-way analysis of variance (Siegel, 1956). Broadbent would predict an insignificant χ^2_r, while Waugh & Norman would predict that the probability of recall should rank in order of the size of L_i. It was found that $\chi^2_r = 15{\cdot}23$, d.f. $= 5$, $P < 0{\cdot}01$. When the estimated value of L_i is compared to the mean probability of recall of items $i = 2$ to 6, there is only one inversion to the rank order predicted by Waugh & Norman. Recall increases as L_i increases, supporting the Waugh & Norman prediction and contradicting both Broadbent and Watkins.

A more direct test of the adequacy of the Waugh & Norman model was then conducted. If Waugh & Norman are correct that both STM and LTM contribute independently to the recall of items in the recency portion of the serial position curve, S_i should be the same for any particular value of i. Watkins and Broadbent, however, believing R_i to be independent of L_i, would predict that the larger the value of L_i, the smaller the value of S_i, when S_i is estimated from the Waugh & Norman formula $S_i = (R_i - L_i) \div (1 - L_i)$.

To test between these predictions, values of S_i were calculated for the items $i = 2$ to 8 for each list used by Murdock, except the ten-item list. This was excluded because examination of the range of the primacy and recency effects in the other lists suggests that these effects overlap for the ten-item list, making an accurate estimate of L_i impossible. Items $i = 7$ and 8 were included since if Waugh & Norman are correct S_i should be constant at any value of i for all items where STM has an effect.

Again, a Friedman two-way analysis of variance was calculated. Here, Watkins and Broadbent predict a significant χ^2_r, with the lists with the greater L_i having the smaller S_i estimates. Waugh & Norman predict an insignificant value for χ^2_r. In fact $\chi^2_r = 6{\cdot}03$, d.f. $= 4$, $0{\cdot}2 > P > 0{\cdot}1$. While this result does not support Broadbent and Watkins it is hardly satisfactory for Waugh & Norman. Examination of the means suggests a tendency to overestimate S_i for the lists with higher L_i values which could result from fewer items being in LTM than is estimated by the Waugh & Norman technique. This possibility, based on the negative recency effect, is a major criticism of the technique and is discussed in detail by Watkins (1974). We do not wish to champion the Waugh & Norman model in its entirety. However, the results that are reported above do seem to support one of its assumptions, namely, that the recency effect represents recall from both STM and LTM and not merely from STM alone. At the least, it has been shown that the variables which affect the main part of the serial position curve also influence the recency effect.

REFERENCES

BROADBENT, D. E. (1971). *Decision and Stress*. New York: Academic.

BROADBENT, D. E. (1975). Memory for events before interruption. *Proceedings of Annual Conference of the British Psychological Society.*

MURDOCK, B. B. Jr. (1962). The serial position effect in free recall. *J. exp. Psychol.* **64**, 482–488.

MURDOCK, B. B. Jr. (1974). *Human Memory: Theory and Data.* Potomac: Lawrence Erlbaum.

NORMAN, D. A. (ed). (1970). *Models of Human Memory*. New York: Academic.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences*. London: McGraw-Hill.

WATKINS, M. J. (1974). Concept and measurement of primary memory. *Psychol. Bull.* **81**, 695–711.

WAUGH, N. C. & NORMAN, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

(*Manuscript received* 18 *April* 1975)

Br. J. Psychol. (1976), 2, 67, *pp.* 165–173
Printed in Great Britain

SEMANTIC CONSTRAINTS AND SENTENCE PERCEPTION

BY STEPHEN E. NEWSTEAD
Psychology Teaching Group, Plymouth Polytechnic

This study investigates semantic constraints in reversible sentences and their relationship to sentence perception. In Expt. I, subjects were given anagrams of reversible sentences to solve, and any preferences for one version of the sentence over the other were taken to indicate the presence of semantic constraints in that sentence. In the second part of the experiment, subjects were presented with these sentences masked with white noise, and there was found to be a relationship between the semantic constraints and the ease of perceiving the sentences. Expts. II and III provide evidence that these constraints derive from a knowledge of the overall relationship expressed in the sentence rather than from the meanings of individual words. Previous experiments have failed to find any effect of semantic constraints in sentence perception; this discrepancy is explained in terms of the strategies subjects might have adopted in the various tasks.

Psycholinguistic research in the past decade has concentrated largely on the problem of syntax. The main concern seems to have been with the 'psychological reality' of the grammatical theories described by Chomsky (1957, 1965). To some extent this emphasis has been justified; although by no means all of Chomsky's ideas have been verified experimentally, one can at the very least say that syntax does play an important role in the perception, memory and understanding of sentences. This is shown for instance by the simple finding that grammatical sequences are easier to process than non-grammatical ones (Epstein, 1961; Miller & Isard, 1963). In addition, there is fairly strong evidence from the 'click' experiments that people actually analyse sentences into their grammatical constituents (Fodor & Bever, 1965; Garrett, Bever & Fodor, 1966).

However, this concentration on syntax has led to a comparative neglect of semantics. Some writers justify this by the claim that syntax is more important than semantics, since the analysis of meaning is subsequent to and dependent upon an analysis of syntax. This 'primacy of syntax' theory seems to have been assumed in most of the early psycholinguistic research, and still has many supporters (e.g. Forster & Ryder, 1971).

Nevertheless, there is an increasing body of evidence that semantic factors are at least as important as syntactic ones. One of the main lines of evidence for this comes from work on semantic constraints, which are restrictions on the ways in which words can meaningfully combine. For instance, there is a semantic constraint inherent in the verb 'to like' which demands that the subject be animate. Thus 'John likes the tree' is a permissible sentence, but not 'The tree likes John'.

The importance of semantic constraints has been shown in work on reversible sentences. These are sentences in which the subject and object can be interchanged and still leave a meaningful sentence, for example 'The boy likes the girl'. Obviously, semantic constraints determine whether or not a sentence is reversible; and Slobin (1966), in an experiment in which subjects had to decide whether various sentences were true or false with respect to a picture, has shown that reversible sentences are more difficult to understand.

Further evidence for the influence of semantic constraints comes from work on self-embedded sentences. Schlesinger (1968) has shown that subjects confronted with such sentences will frequently give the most reasonable interpretation in terms of the semantic constraints, ignoring completely the actual syntax of the sentence. Thus subjects had little difficulty in assigning a meaningful interpretation to the sentence 'This is the hole, that the rat, which our cat, whom the dog bit, made, caught'; it seems obvious that the rat made the hole, the cat caught the rat, and so on. However, a close analysis of the sentence reveals that this semantically reasonable interpretation is in fact grammatically incorrect; what the sentence actually says is that the rat caught the hole, the cat made the rat, and the dog bit the cat. Similar findings to Schlesinger's have been reported by Stolz (1967) and Newstead (1972).

Semantic constraints of a different kind were investigated in an experiment by Herriot (1969). Herriot talks not of semantic constraints but of 'pragmatic expectancies', which involve constraints deriving from our knowledge of the world rather than the actual meanings of the words. To illustrate, Herriot used reversible sentences such as 'The doctor treated the patient'. Although this sentence is theoretically reversible, our knowledge of the world tells us that it is unlikely to be reversed, since patients do not normally treat doctors. Herriot's experiment, which involved extraction of the logical subject and logical object of each sentence, showed that such sentences were easier to analyse in their normal direction. There is no useful purpose to be served here by distinguishing these pragmatic expectancies and other semantic constraints, and so this experiment can be taken as further evidence for the importance of semantic constraints in sentence processing.

More recently, Forster & Olbrei (1973) have criticized this work on semantic constraints. They claim that both Slobin's and Herriot's experiments involved somewhat artificial tasks, and this might explain the obtained results. For instance, in Slobin's experiment, the pictures could have had more effect on the difficulty of the reversible sentences than the sentences themselves. With a reversible sentence such as 'The boy is hitting the girl', subjects would have had to examine the picture closely to see if it was the boy or the girl who was performing the action, and to whom; but with a non-reversible sentence such as 'The girl is watering the flowers', a cursory examination may be sufficient to determine whether or not there is a girl performing the action described. Forster & Olbrei carried out two experiments of their own which indicated that reversible sentences were no more difficult than non-reversible sentences. The first experiment involved judgements of the acceptability in English of various sentences; the second involved reporting back as many words as possible from sentences which were presented rapidly on a screen, one word at a time.

There is not necessarily a contradiction between the results of these experiments and those obtained by Slobin and Herriot. It is possible that reversible sentences are more difficult than non-reversible sentences only when the sentence needs to be analysed to the deep structural or semantic level. A fairly superficial examination of surface structure was all that was required in both of Forster & Olbrei's experiments, but Slobin's and Herriot's experiments required an understanding of the meaning of the sentences. Significantly, when Forster & Olbrei used another task which required the semantic processing of sentences, reversible sentences proved more difficult, even though the effect was small. Unfortunately, this initially plausible explanation is not consistent with the findings reported in the present paper, a point which will be considered further in the general discussion.

The present research is a more precise analysis of the way in which semantic constraints affect the perception of reversible sentences. Previous research has simply compared the difficulty of reversible and non-reversible sentences; in the present research it is assumed that even within reversible sentences there may exist semantic constraints which make some sentences more reversible than others. It has already been mentioned that some reversible sentences contain quite powerful semantic constraints which make them effectively non-reversible (for instance, Herriot's 'The doctor treated the patient'). In other reversible sentences, the constraints seem less powerful, if they exist at all. The sentence 'The boy is pushing the girl' seems quite natural in both directions; nevertheless it is still possible that quite subtle constraints exist even here. For example, our knowledge of sex roles might lead us to expect that males are more likely to do the pushing than females.

In Expt. I, a technique is described for measuring the semantic constraints in reversible sentences. The measure was termed the 'degree of reversibility', and in the second part of the experiment this was correlated with the ease of perceiving the sentence in white noise. Expts. II and III investigate in more detail the nature of the semantic constraints involved in sentences such as these.

EXPERIMENT I

Method

The measurement of semantic constraints

A questionnaire was compiled containing anagrams of the 20 sentences shown in Table 1. There were in fact two versions of the questionnaire, the only difference between them being that the subject and object nouns were interchanged; this was to control for any effects of anagram word-order on the solutions given. The questionnaire was given to 94 male and female subjects, each version being given to half the subjects: 74 of the subjects were secondary-school children from Greenwood Dale Bilateral School, Nottingham, while the other 20 were students at Nottingham University.

An analysis of variance by age, sex and questionnaire form revealed no significant differences, and so the data from all subjects were pooled. The results are summarized in Table 1, in terms of the 'degree of reversibility' of each sentence. This was defined as the ratio between the number of times the non-preferred version was given and the number of times the preferred version was given, expressed as a percentage. In Table I each sentence is given in what turned out to be its preferred direction. Obviously, the degree of reversibility is a measure of how powerful the semantic constraints in the sentence are. With strong constraints, the sentence will be almost totally non-reversible (e.g. 'The dog is chasing the cat'); while sentences containing fewer constraints will be almost totally reversible (e.g. 'Daughters confide in their mothers'). The main part of the experiment attempted to correlate degree of reversibility with ease of sentence perception.

Material

Exactly the same sentences as given in Table 1 were used. Using normal intonation, the investigator recorded 32 tapes of the 20 sentences to be presented to each subject; all of these tapes were then masked with white noise. Ten of the sentences on each tape were read out in their preferred direction and ten in their non-preferred direction. Alternate tapes used the same sentences in the same order as the previous tape, but with the preferred and non-preferred versions interchanged. Within these restrictions, the order of the sentences on each tape was randomized.

Table 1. *Sentences used and their degree of reversibility*

Sentence	Degree of reversibility (%)		Significance level of difference between Expts. I and III
	Expt. I	Expt. III	
1 The dog is chasing the cat	0	55	< 0·001
2 The lady described the visitor	16	45	n.s.
3 Teachers like their pupils	52	−11	< 0·001
4 The Americans imitate the English	63	−78	n.s.
5 The secretary is following the manager	19	33	n.s.
6 Parents understand their children	31	52	n.s.
7 The lawyer talked to the judge	48	−23	< 0·001
8 Policemen get angry with criminals	32	−63	< 0·001
9 Adults trust teenagers	66	68	n.s.
10 The customer insulted the barman	71	78	n.s.
11 The bishop believed the vicar	52	−82	n.s.
12 Daughters confide in their mothers	91	−19	< 0·01
13 The soldier is kissing the nurse	33	23	n.s.
14 Husbands love their wives	74	−28	< 0·01
15 The boy is pushing the girl	15	−78	< 0·001
16 The French hate the Germans	57	−72	n.s.
17 All people like animals	16	33	n.s.
18 Men work harder than women	34	−60	< 0·01
19 The councillor lied to the mayor	22	78	< 0·01
20 The man is hitting the boy	4	19	n.s.

Subjects and procedure

Subjects were 32 paid volunteers of both sexes, all of whom were students at Nottingham University. None of them had participated in the first part of the experiment. They listened to the 20 masked sentences and after each one wrote down what they thought the sentence was. Four trial sentences were presented to accustom subjects to the task.

Results

The sentences produced were scored either correct or incorrect ignoring spelling mistakes and changes of tense and number. To test if subjects found the sentences presented in their preferred direction easier to perceive correctly than those presented in their non-preferred direction, a Wilcoxon matched-pairs signed ranks test was carried out. There was a significant difference, T (28) = 128, $P < 0{\cdot}05$, one-tailed test. To determine whether the degree of reversibility was related to the accuracy of perception, a Spearman rank-order correlation was carried out. The two factors to be correlated were degree of reversibility (see Table 1) and the difference between the number of correct answers given to the preferred and non-preferred versions of each sentence. This yielded a correlation of 0·43, which is statistically significant ($P < 0{\cdot}05$, one-tailed).

DISCUSSION

This experiment leaves little doubt as to the importance of semantic constraints in sentence perception. It has been shown that sentences are easier to perceive when they conform to the rather subtle constraints revealed in the preliminary questionnaire. Further, the strength of these constraints has also been shown to be related to ease of perception. Hence this experiment confirms the previous research which has shown the importance of semantic constraints; the main difference is that the constraints used in this experiment were rather more subtle than those used by previous experimenters.

It is of interest to speculate as to the exact nature of the constraints involved in the present experiment. There appear to be three possibilities; that they derive from the formal denotational meanings of the individual words; that they derive from the connotational meanings of the words; or that they involve knowledge of the overall relationship expressed in the sentence. It seems unlikely that the formal meanings of words are important: the knowledge that 'boy' is human, male, not-adult, etc., does not lead to the expectation that boys should push girls. It is, however, possible that the connotational meanings are relevant. Johnson (1967) has shown that words which are the subjects of sentences tend to be rated as more active on the semantic differential than do the objects. Hence the preferences in the anagram study may have resulted simply from subjects putting the more active nouns in the subject position.

Expt. II investigates this possibility by examining the ratings on a semantic differential of the nouns in the sentences used in Expt. I. If the prediction taken from Johnson's experiment is correct, then the subject nouns should be rated consistently more active than the object nouns. There is also the possibility that the subject nouns might have been chosen because they were more 'appropriate' to the verb (for instance an active verb might require an active subject, a passive verb a passive subject). To enable this to be tested the verbs used in Expt. I were also rated on the semantic differential.

EXPERIMENT II

Method

Material

A booklet was prepared containing the 20 subject nouns, 20 object nouns and 20 verbs from the sentences used in the previous experiment (see Table 1). The words were presented in different random order in each booklet. Beneath each word there were seven seven-point bipolar scales, labelled strong–weak, active–passive, energetic–inert, subject–object, aggressive–defensive, potent—impotent and fast–slow. All of these are related to the activity and potency dimensions of the semantic differential.

Subjects and procedure

The booklet was distributed to 20 subjects, male and female undergraduate volunteers from Bolton Institute of Technology. They were asked to fill in the ratings in their own time, but without consulting each other. It was stressed that their first impressions were more important than a long-studied opinion.

Results and discussion

Results are summarized in Table 2. All significance levels are based on a two-tailed Wilcoxon matched-pairs signed ranks tests. For the purposes of this analysis, the average rating of each noun and verb by each subject was used. As can be seen from this table, in 12 of the 20 sentences there was a significant difference in the activity/potency ratings of the subject and object nouns. However, this difference was in the predicted direction for only six of the sentences; for the remaining six, the object nouns were rated as more active and potent than the subjects! Overall there was no significant difference in the ratings of subject and object nouns; if

Table 2. *Semantic differential ratings in Expt. II*

Sentence	Comparison of activity/potency of subject and object nouns	Appropriateness of subject noun to verb (compared to object noun)
1	$<0{\cdot}01$	n.s.
2	n.s.	n.s.
3	$<0{\cdot}01$	n.s.
4	$<0{\cdot}01$	n.s.
5	$<0{\cdot}01$*	n.s.
6	n.s.	n.s.
7	n.s.	n.s.
8	n.s.	n.s.
9	$<0{\cdot}01$*	n.s.
10	$<0{\cdot}01$*	n.s.
11	$<0{\cdot}01$*	n.s.
12	$<0{\cdot}01$*	n.s.
13	n.s.	n.s.
14	$<0{\cdot}02$	n.s.
15	$<0{\cdot}01$	n.s.
16	$<0{\cdot}01$*	$<0{\cdot}01$
17	n.s.	n.s.
18	$<0{\cdot}01$	n.s.
19	n.s.	n.s.
20	n.s.	n.s.

An asterisk indicates that the result went *against* the prediction.

anything, there was a slight tendency for subject nouns to be rated *less* active and potent than object nouns.

Similarly, there was no systematic relationship between the activity/potency of the verb and the kind of subject noun it required. For only one of the 20 sentences was the rating of the subject noun significantly closer to the rating of the verb than was the object noun. Overall, there was no significant tendency for the subject nouns to be more appropriate to the verb than the object nouns. ($n = 19$, $x = 9$, $P = 0{\cdot}5$, sign test.)

It is obvious from these results that the semantic constraints present in these reversible sentences cannot be easily explained in terms of the connotational meanings of the words. There was no significant tendency to put active and potent words in the subject position; nor was there a tendency to put active subjects with active verbs (and passive subjects with passive verbs). It is not impossible that the constraints can be explained in terms of some other aspect of connotational meaning, although this does seem intuitively unlikely. It seems far more probable that the constraints derive from our knowledge of the relationship expressed in the sentence, rather than from individual word meanings. If this is so, then one would expect that the preferences shown by subjects in solving anagrams of reversible sentences would be changed if the relationship expressed in the sentence was changed. This prediction is investigated in Expt. III.

EXPERIMENT III

Method

Material

This experiment involved measuring preferences for the direction of reversible sentences using the anagram technique described in Expt. I. The sentences used were those given in Table I, but with the verbs changed. This was achieved simply by moving the verbs up one sentence: thus the verb 'describe' went from sentence 2 to sentence 1, the verb 'like' went from sentence 3 to sentence 2, and so on. One or two changes of tense and number were required to make the sentences grammatically acceptable after these changes. A questionnaire containing anagrams of the 20 sentences produced was prepared, exactly as in Expt. I.

Subjects and procedure

The questionnaire was distributed to 32 male and female psychology undergraduates at Plymouth Polytechnic.

Results and discussion

For ease of presentation, the results are shown in the second part of Table 1. The degree of reversibility of each sentence is given, a negative number indicating that the sentence was preferred in the opposite direction to Expt. I. For 10 of the 20 sentences, the preferences went in exactly the opposite direction to Expt. I. Further, for 9 of the 20 sentences there was a significant difference in the preferences shown in the two experiments. Overall, there was an insignificant negative correlation (– 0·22, Spearman rank-order correlation) between the degrees of reversibility found in the two experiments.

These results suggest quite clearly that the anagram preferences are readily modified if the relationship expressed in the sentence is changed. In other words, the semantic constraints in these reversible sentences seem to stem from our knowledge of the relationship described in those sentences.

GENERAL DISCUSSION

This series of experiments has shown three things: firstly, that there exist semantic constraints even in apparently reversible sentences; secondly, that the power of these constraints is related to ease of sentence perception; and thirdly, that the constraints derive from the overall relationship described in the sentence.

The finding that these semantic constraints affect performance in a perceptual task (perception in white noise) is of some interest. It would seem to call into question the explanation given in the introduction of the discrepancy between the results of Slobin (1966) and Herriot (1969) and Forster & Olbrei (1973). There it was suggested that Forster & Olbrei's results may be explained by the fact that they used a perceptual task, while the other researchers used tasks involving the analysis of deep structure and meaning. The present experiment indicates that semantic constraints can prove effective in at least one kind of perceptual task.

It is difficult to find any obvious explanation as to why semantic constraints should operate in the present perceptual task, but not in the tasks used by Forster & Olbrei. One possible explanation is in terms of the strategies that subjects adopted in the experiments. Subjects who are faced with the task of perceiving sentences in white noise might 'reconstruct' a sentence around the few fragments of the sentence that have been perceived. If so, they might well bring in their knowledge of grammar and meaning to facilitate this reconstruction. There was some evidence that subjects were doing just this in Expt. I, since more than 40 per cent of the errors made involved giving full, meaningful sentences, but not the one that had been presented; it seems that subjects failed to perceive all the words in the sentence and so constructed a meaningful sentence around what they had perceived. The tasks used by Forster & Olbrei – judging the acceptability in English of sentences, and reporting back words from sentences which were presented on a screen – might have led to rather different strategies not involving perceptual reconstruction. However, some writers (e.g. Neisser, 1967) have claimed that all perception is reconstructive in nature, and if this is so there should be no such difference in strategy between the various kinds of perceptual tasks. Whatever the explanation might be, the finding remains that very subtle semantic constraints can affect performance in at least some perceptual tasks.

Since the semantic constraints studied seem to derive from a knowledge of the overall relationship described in the sentence, this poses questions about the precise nature of this knowledge. Is this knowledge deep seated and permanent, or is it more transient and modifiable by experience? One way of investigating this question is to study whether it is possible to change the semantic constraints that exist in reversible sentences. Herriot *et al.* (1974) have turned their attention to this problem. Using the anagram technique described in this paper, they have shown that the solution given can be quite easily modified. People normally give the version 'The soldier kissed the nurse', but if they have just read a passage about nurses kissing soldiers then they tend to give the opposite version. Hence these semantic constraints are not hard and fast rules about language use; and yet they do affect sentence perception, as was shown in Expt. I.

The study by Herriot *et al.* (1974) suggests that the preferences shown in the anagram solution task reflect the subject's expectancies at that particular point in time. Hence it might be possible to develop the technique as a means of measuring a person's perception of and attitudes towards the world. For instance, it might be possible to develop an MCP (male chauvinist pig) scale, predicting that chauvinists would give the version 'Men work harder than women', 'Boys are more intelligent than girls', and so on. If such items were embedded in a large questionnaire, it would not be transparently obvious exactly what was being measured. One advantage of such a technique of attitude measurement is that, as illustrated in the study by Herriot *et al.*, it is quite sensitive to sudden changes.

This research was carried out at the University of Nottingham, and was supported by a grant from the Medical Research Council. I would like to thank Dr Jonathan Evans for his critical reading of an earlier draft of this paper.

References

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass: M.I.T. Press.

Epstein, W. A. (1961). The influence of grammatical structure on learning. *Am. J. Psychol.* 74, 80–85.

Fodor, J. A. & Bever, T. G. (1965). The psychological reality of linguistic segments. *J. verb. Learn. verb. Behav.* 4, 414–420.

Forster, K. I. & Olbrei, I. (1973). Semantic heuristics and syntactic analysis. *Cognition* **2**, 319–347.

Forster, K. I. & Ryder, L. A. (1971). Perceiving the structure and meaning of sentences. *J. verb. Learn. verb. Behav.* **10**, 285–296.

Garrett, M., Bever, T. G. & Fodor, J. A. (1966). The active use of grammar in speech perception. *Percept. Psychophys.* **1**, 30–32.

Herriot, P. (1969). The comprehension of active and passive sentences as a function of pragmatic expectancies. *J. verb. Learn. verb. Behav.* **8**, 166–169.

Herriot, P., Craggs, K., Fawcett, T., Stevenson, J. & Still, B. (1974). Ambiguous anagrams. *Br. J. soc. clin. Psychol.* **13**, 331–332.

Johnson, M. G. (1967). Syntactic position and rated meaning. *J. verb. Learn. verb. Behav.* **6**, 240–246.

Miller, G. A. & Isard, S. (1963). Some perceptual consequences of linguistic rules. *J. verb. Learn. verb. Behav.* **2**, 217–228.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

Newstead, S. E. (1972). Semantic factors in sentence processing. Unpublished PhD thesis, University of Nottingham.

Schlesinger, I. M. (1968). *Sentence Structure and the Reading Process*. The Hague: Mouton.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *J. verb. Learn. verb. Behav.* **5**, 219–227.

Stolz, W. S. (1967). A study of the ability to decode grammatically novel sentences. *J. verb. Learn. verb. Behav.* **9**, 867–873.

(*Manuscript received* 11 *October* 1974; *revised manuscript received* 24 *April* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 175–180
Printed in Great Britain

EFFECTS OF NOVELTY AND ODDITY ON VISUAL SELECTIVE ATTENTION

By D. E. BERLYNE and JOYCE DITKOFSKY
Department of Psychology, University of Toronto, Toronto, Canada

Effects of novelty and of oddity on selective attention were investigated through tachistoscopic recognition. When a pre-exposure field consisted of dots of one colour, letters of a different colour were more likely to be identified than letters of the same colour. In displays of 10 letters, two letters differing in colour from the remainder were more likely to be identified. Tachistoscopic exposure precluded determination of responses by eye movements, and there were controls for chromatic adaptation and other factors that might have biased the results.

The term 'attention' has been used in psychological literature to denote a number of distinct phenomena (Berlyne, 1969, 1970, 1974). One of them is the selection of stimulus events located in a particular region of the stimulus field (or stimulating particular receptors) to exert preponderant control over behaviour.

A technique for the study of selective attention in this sense was used in several earlier experiments. The subject goes through a succession of free-choice trials, on each of which two or more stimuli appear in locations corresponding to distinct, mutually exclusive key-pressing responses. He is free to respond to any one of the stimuli. The free-choice trials are interspersed with forced-choice trials, on which only one stimulus occurs and the response corresponding to its location is to be performed.

With the help of this technique, the probability of responding to a particular stimulus, rather than to others that are competing with it for attention, has been shown to depend on colour, intensity and related variables (Berlyne, 1950, 1970; McDonnell, 1970). Other experiments took subjects first through a habituation phase of trials with stimuli of one colour and shape and then through a test phase, in which some stimuli were like those of the habituation phase but others were of a different colour or shape. There was a significantly greater tendency to respond to the latter, which means that novel stimuli tended to attract attention (Berlyne, 1951). Subsequent experiments confirmed this effect of novelty with controls for additional factors. The effect appeared when stimuli were exposed tachistoscopically for 0·1 sec, excluding the possibility that the response might be determined by the direction of eye movements (McDonnell, 1968; see Berlyne, 1970). It also appeared when stimuli presented during the habituation phase and the test phase occupied different locations in the visual field and thus stimulated different retinal points, excluding the possibility that chromatic adaptation may have been responsible (Berlyne, 1971).

The experiments to be reported were also concerned with selective attention but used a somewhat different technique, namely tachistoscopic recognition of letters of the alphabet. When a stimulus is successfully competing with others for attention, i.e. for control over behaviour, information from that stimulus is occupying the limited channel capacity of the subject's nervous system and being transmitted to the motor system, whereas information originating in the other stimuli is blocked. In the experiments using the free-choice technique, only information about the location of the stimulus had to be transmitted. When, however, letters have to be named, other information, i.e. about shape, has to be taken in and allowed to determine the response.

Two of the experiments were intended to investigate further the effects of novelty on attention. First, it was hoped to ascertain whether information about other properties than location is more readily absorbed from novel stimuli than from others that are simultaneously present. Secondly, it was hoped that further light would be shed on the role of the task performed during habituation. In most of the free-choice experiments already mentioned, the same task, namely selecting a button to press, was carried out during habituation and test phases. The novelty effect also appeared when a different task, i.e. naming the location of a stimulus, was used in the habituation phase and button-pressing in the test phase or vice versa (Berlyne, 1971). But both of these tasks involved choice. The effect failed to appear when a task not involving choice, i.e. guessing the number of stimuli that would appear on the next trial, characterized the habituation phase (Berlyne, 1957). It was conjectured that novelty might affect attention even when no overt response occurred during habituation, provided that the novel stimuli appeared immediately after habituation stimuli, thus obviating the weakening of the novelty effect that could well result from an inter-trial interval of a few seconds. A technique fulfilling this requirement was contrived for the present project.

The other two experiments concentrated on a variable that had not been studied in earlier experiments, namely oddity, i.e. the property of being in a minority, of differing in some salient respect from most of the accompanying stimuli. There is a sense in which oddity is a spatial equivalent of novelty. A novel stimulus is one that differs from most of the stimuli that have been experienced in the recent past, whereas an odd stimulus is one that differs from others that are experienced at the same time.

EXPERIMENTS IN AND IO

Subjects

Twenty-four undergraduates taking introductory psychology courses participated in three experiments consecutively. The first was an experiment that will not be reported in this article but also involved tachistoscopic recognition of letters. Then followed Experiment IN (Novelty) and, after it, Experiment IO (Oddity). There was a 5 min rest period between experiments, at the end of which the instructions for the next experiment were given.

Apparatus

A Gerbrands Harvard-type two-channel tachistoscope, connected with a Hunter timer, exposed a test field for 0·1 sec when manually triggered. Between exposures of the test field, a pre-exposure field was visible. Both fields were seen 56·5 cm from the subject's eyes and occupied a 7·1 × 7·1 cm square. Both were illuminated with independent cold white fluorescent lamps (Sylvania, P-4 daylight).

Experiment IN

Stimulus material. Two pre-exposure cards and six test cards were prepared. All eight cards had a black cross in the centre to indicate the fixation point. The pre-exposure cards bore eight dots of 0·6 cm radius, equally spaced around the circumference of a 3 cm radius circle centred on the cross. On one pre-exposure card, the dots were red (Stabilo-Pen 68/48), and, on the

Table 1. *Mean numbers of letters recognized*

				Pairs clockwise from odd pair				
Expt.	Novel	Non-novel	Expt.	1 (odd)	2 (adj.)	3 (non-adj.)	4 (non-adj.)	5 (adj.)
IN	8·2	6·3	IO	4·8	3·5	4·3	3·6	3·9
IIN	8·2	6·0	IIO	4·1	3·6	3·3	2·9	3·3

other card, they were green (Stabilo-Pen 68/43). The six test cards bore four red and four green letters (drawn with Letraset Instant Lettering, No. 2153, 48 pt Helvetica Medium Outline). Each letter was 1·2 cm in height and 1 cm in width and appeared midway between locations occupied by two dots on the pre-exposure cards. This was to make sure that letters and dots did not stimulate the same points on the retina, thus controlling for chromatic adaptation. Red and green letters alternated around the circle. The letters were randomly selected from the alphabet with two restrictions: (*a*) no letter appeared twice on any one card, and (*b*) over all six test cards, every letter in the alphabet appeared with approximately equal frequency.

Procedure. The subjects were randomly divided into two groups of 12, R and G. The pre-exposure field consisted of the card bearing the red dot for group R (for whom the green letters on the test cards were consequently the novel stimuli) and the card bearing the green dot for group G (for whom the red letters were the novel stimuli). Subjects were simply told that they would see a series of cards bearing letters and were to report as many letters as they could as soon as a test card became visible. The experimenter said 'Ready!' shortly before presenting each test card, and the subjects were told to keep their eyes on the black cross on receiving this signal and while seeing the test card. The test cards were presented in the same randomly selected order for every subject.

Results. Mean numbers of novel and non-novel letters recognized appear in the first row of Table 1. Since there were six test cards, each of which bore four novel and four non-novel letters, each of these means may be compared with a maximum possible score of 24. Analysis of variance confirms that significantly more novel stimuli were identified ($F = 15{\cdot}25$; d.f. = 1,22; $P < 0{\cdot}001$).

Experiment IO

Stimulus material. Two sets of five stimulus cards were prepared. Each had a black cross in the centre, and the lettering had the same specifications as the test cards of Experiment IN. Ten letters were equally spaced around a 4 cm radius circle centred on a black cross. On each card, there were two adjacent letters of one colour (odd) and eight letters of another colour (non-odd). A different pair of colours was used for each of the five cards in a set, and each card belonging to one set had a corresponding card in the other set. The same two colours were used for both, but the colour used for the two odd letters in one card were used for the eight non-odd letters on its counterpart, and vice versa. The same letters, appearing in the same locations and in the same colour, were odd in one set and non-odd in the other set. Furthermore, in each set of five cards, every position around the circle was occupied by an odd letter on one of the cards and by a non-odd letter on the remaining cards.

Procedure. The subjects were randomly divided into two groups of 12, each consisting of six subjects from the R group and six from the G group of Experiment IN. Each group had one of the two sets of test cards. The pre-exposure field consisted of a blank white card with a black cross. The test cards were presented in the same random order to all the subjects of one group, and the corresponding test cards of the other set were presented to the other group in the equivalent order. Consequently, the two letters that were odd for one group of subjects on a particular trial were non-odd for the other group but were nevertheless seen on the same trial in the same locations and in the same colour.

Results. The mean numbers of odd letters identified over all five test cards was computed. The corresponding means were also computed for the remaining four pairs of letters. These data are displayed in the first row of Table 1, proceeding clockwise from the odd pair. Since there were five test cards, the maximum possible score is 10 in each case.

Two planned orthogonal contrasts were examined. First, the mean for the odd pair significantly exceeded the mean for the non-odd pairs ($F = 10{\cdot}29$; d.f. = 1,88; $P < 0{\cdot}001$). Secondly, among the non-odd pairs, those that were adjacent to the odd pair (pairs 2 and 5 in Table 1) were compared with non-adjacent pairs (3 and 4). This was to ascertain whether there was a spread-of-effect or spatial-gradient phenomenon. The F value for this contrast did not, however, approach significance.

EXPERIMENTS IIO AND IIN

The aim of the next two experiments was to find out whether the results of the first two would be confirmed. In particular, it will be recalled that the previous subjects went through Experiment IN just after an unreported experiment and through Experiment IO after Experiment IN. This time the unreported experiment was omitted, and the order in which the experiments were carried out was reversed, thus controlling for the possibility that the results of Experiments IN and IIO might in some way have depended on the experiences that preceded them.

Subjects

Twenty-four additional subjects were recruited through posters from undergraduates attending the university summer session. They were paid for their participation. These subjects went through exactly the same procedures as those of Experiments IN and IO but in the opposite order. In other words, they had the oddity experiment (IIO) first and the novelty experiment (IIN) second.

Results

The means are shown in the second row of Table 1. They closely resembled those of the previous experiments. In Experiment IIN, significantly more novel than non-novel letters were once again recognized ($F = 12{\cdot}69$; d.f. = 1,22; $P < 0{\cdot}001$). In Experiment IIO, odd letters were more likely to be identified than non-odd letters ($F = 6{\cdot}54$; d.f. = 1,88; $P < 0{\cdot}02$), and there was again no significant difference between non-odd letters that were adjacent and non-adjacent to the odd letters.

DISCUSSION

The findings are therefore reasonably clear-cut and evidently reproducible. Novel letters, i.e. letters whose colour differs from one to which the subject has just been exposed, are more likely to be recognized than non-novel letters. The use of the tachistoscopic presentation excludes the possibility that this phenomenon may occur because novel letters attract eye movements and fixated stimuli receive priority in perceptual processing. The presentation of pre-exposure and test stimuli in different portions of the visual field likewise excludes the possibility that the effect may be due to peripheral chromatic adaptation.

The result of Experiments IO and IIO indicate that odd stimuli, i.e. those that contrast with most of the stimuli in the visual field, have an advantage in the competition for selective attention. There is no sign that this advantage spreads to adjacent stimuli.

One earlier experimenter who obtained some evidence for the power of oddity to attract attention is Wilcocks (1928). He presented letters for 15 min and later tested for recall. If one letter was moving while the rest were stationary or vice versa, the exceptional letter was more likely to be remembered. There are, however, several obvious differences in procedure between the present experiments on oddity and the one by Wilcocks.

The demonstrated tendency for odd stimuli to capture attention is reminiscent of the von Restorff phenomenon (von Restorff, 1933), the relative ease of recalling a verbal item that contrasts in some way with most of the series to which it belongs. This phenomenon involves, however, attention in remembering rather than attention in performance (Berlyne, 1960), and it is by no means clear how far these two kinds of selective attention depend on the same underlying processes. Furthermore, later investigators (e.g. Green, 1956; see Wallace, 1965) have raised doubts about whether oddity (or 'isolation', as it is often called) is the crucial factor. Our results may also seem related to another phenomenon that figured prominently in the writings of the Gestalt school, namely Rubin's (1915) figure–ground phenomenon. There has been some evidence (e.g. Graham, 1929) that units occupying a relatively small part of the visual field are more likely to appear as figure, and there have also been suggestions that the figure is the part of a display that receives more attention. However, Rubin and the Gestalt school repudiated this suggestion (see Berlyne, 1974). All in all, therefore, the connexions between our findings with regard to oddity and these other phenomena are obscure.

On the other hand, the connexion between oddity and selective attention may make sense in terms of biological adaptation. Objects or events that contrast with most of the stimulus field seem more likely than others to betoken conditions of biological significance that call for action, so that concentration on them may be appropriate. Sources of mildly exceptional or isolated stimulation may tend to evoke approach, whereas sharper contrast leads to withdrawal and signs of distress (Schneirla, 1959; Andrew, 1964; Berlyne, 1967). In either case, attention and investigation would appear to be indispensable preliminaries.

This investigation was supported by research grant A-73 from the National Research Council of Canada. Experiments IN and IO were reported in a thesis submitted by the junior author in accordance with the requirements for the MA degree of the University of Toronto.

REFERENCES

ANDREW, R. J. (1964). Vocalization in chicks and the concept of 'stimulus contrast'. *Anim. Behav.* **12**, 64–76.

BERLYNE, D. E. (1950). Stimulus intensity and attention in relation to learning theory. *Q. Jl exp. Psychol.* **2**, 71–75.

BERLYNE, D. E. (1951). Attention to change. *Br. J. Psychol.* **42**, 269–275.

BERLYNE, D. E. (1957). Attention to change, conditioned inhibition ($_{S}I_{R}$) and stimulus satiation. *Br. J. Psychol.* **48**, 138–140.

BERLYNE, D. E. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

BERLYNE, D. E. (1967). Arousal and reinforcement. In D. Levine (ed.), *Nebraska Symposium on Motivation, 1967*. Lincoln, Neb.: University of Nebraska Press.

BERLYNE, D. E. (1969). The development of the concept of attention in psychology. In C. R. Evans & T. Mulholland (eds), *Attention in Neurophysiology*. London: Butterworth.

BERLYNE, D. E. (1970). Attention as a problem in behaviour theory. In D. Mostofsky (ed.), *Attention: Contemporary Theory and Analysis*. New York: Appleton-Century-Crofts.

BERLYNE, D. E. (1971). Novelty and attention: controls for retinal adaptation and for stimulus response specificity. *Psychon. Sci.* **25**, 349–351.

BERLYNE, D. E. (1974). Attention. In E. C. Carterette & M. P. Friedman (eds), *Handbook of Perception*, vol. 1. New York: Academic Press.

GRAHAM, C. M. (1929). Area, colour and brightness difference in a reversible configuration. *J. gen. Psychol.* **2**, 470–481.

Br. J. Psychol. (1976), 67, 2, *pp.* 181–189
Printed in Great Britain

CONDITIONING AFTERIMAGES: A PROCEDURE MINIMIZING THE EXTINCTION EFFECT OF NORMAL TEST TRIALS

BY PETER DAVIES

Department of Psychology, University of Bradford

Six subjects were trained on a conditioning schedule of ten trials a day for 25 days: a further 18 CS–US pairings were presented on 4 subsequent days. The tone employed as the CS was put on 30 sec before the presentation of a briefly illuminated visual target (the US) and maintained until terminated by the subject when his afterimages disappeared. This procedure allowed continuous monitoring of the evolution of the conditioned response (conditioned afterimages): directly for CRs with a latency of less than 30 sec which could occur on every training trial before the US was presented; indirectly from the changes in the duration of afterimagery following the presentation of the US. As every trial yielded some evidence about the evolution of the CR, unreinforced test trials were not necessary after every block of training trials so minimizing the extinction effect attributable to the presentation of an unreinforced CS.

In the course of the experiment the mean duration of afterimagery increased almost fivefold with one subject showing a tenfold increase. Five of the subjects experienced visual images in response to the tone alone and were judged to be conditioned. The data are discussed in relation to the evolving levels of conditioning identified by Bzhalava (1958, 1965).

Razran (1971) summarized work by Bzhalava (1958, 1965) in which the test trials were combined with the training trials. The conditioned stimulus, a tone, was presented for 30 sec of isolated action before the unconditioned stimulus (2 sec of illumination of a target) was presented. The tone was continued after the presentation of the US until the subject reported the termination of afterimagery.

Bzhalava's results suggest that this method has some advantages over the use of brief tone presentations as a CS and the use of normal unreinforced test trials employed previously (Davies, 1974*a*, *b*). The subject is tested on every trial during the first 30 sec, and is then reinforced by the presentation of the US on every trial. Moreover, the conditioned stimulus is paired not only with the unconditioned stimulus but also with the unconditioned responses which thus acquire the status of secondary unconditioned stimuli. Indirect evidence of the evolution of the conditioned response may be obtained from the duration of each series of afterimages, and a large number of test trials, equal to the number of training trials, may be employed without the test trials having an extinction effect. It was, therefore, decided to employ Bzhalava's procedure in the experiment to be reported here.

SUBJECTS

Six subjects were used, four females and two males. Four of the subjects were drawn from the 20–25 age group, two of the females from the 35–40 age group. All were paid for their services.

APPARATUS

The subject was seated in a totally dark room in front of an apparatus designed to present a brief light flash to his left eye. The flash illuminated a patterned target (circle with inscribed cross) which was the unconditioned stimulus giving rise to unconditioned visual afterimages. The subject wore headphones to enable him to hear verbal instructions and the 1000 Hz tone (the CS), and communicated with the experimenter by talking into a microphone. A low-intensity table lamp, normally off, was provided for the subject's use together with prepared drawing books and coloured pens. Each page of the drawing books had eight outline circles drawn on it to indicate the normal size of afterimage, the pens were all clearly labelled with their colour names as colours could not be distinguished reliably owing to the dim light employed.

The experimenter's room was adjacent to the subject's room and contained a variety of recording apparatus, timers, and stimulus control equipment. A more complete description of the experimental arrangements is to be found elsewhere (Davies, 1974*a*).

METHOD

The subjects were tested individually. Their task was to sit in the totally dark room and report all their visual experiences by describing them as they happened, to depress and hold a response key whenever they were experiencing a visual image of any kind, to stop the tone when the images finally ceased, and to draw one sequence of images when asked to do so. Hearing the tone was the cue to look into the apparatus where 30 sec after the tone onset their left eye was stimulated briefly by a single 1 msec flash of the target. After one trial each day, determined in advance of the session in order to avoid selecting trials on the basis of the report content, the subject was asked to draw the images reported on that trial. This was done immediately after the end of the report with the subject's recall being assisted by replaying their own tape-recorded verbal description. In order not to lose the dark adaptation of the left eye they covered their left eye with an opaque close-fitting eye shield before the table lamp was switched on to enable them to see to draw.

On the first day each subject was given typewritten instructions to 'report anything that you think you are seeing whenever it should occur' and told how to do so as outlined above. Instructions for tone termination and the drawing procedure were also explained to the subject. In order to check that the subject understood the instructions, they were given one practice trial while still light adapted and in the presence of the experimenter. Real tone was not employed, they were told 'Tone now on' as the signal to look into the apparatus. The procedure for drawing was practised after this trial.

The subject was then left alone in the now darkened room for 5 min before commencing ten trials of flash only at a mean intertrial interval of 3·5 min. The resulting mean afterimage duration was taken as the base level for unconditioned afterimages.

On the following day conditioning was commenced at the rate of ten trials per day (mean i.t.i., 3·5 min; range 3–4 min) and this schedule was maintained for 25 days. The days were consecutive apart from minor interruptions for all subjects except S. 2, who had her conditioning interupted by three months of illness.

Each session started with 5 min dark adaptation before the first trial. The normal trial was 30 sec of 1000 Hz tone of moderate intensity, a brief flash of the target, and continuation of the tone until the subject terminated it when afterimagery ceased. Subjects were instructed to let the tone run on for what they estimated to be 10 sec after the final afterimage had disappeared in case another should appear. Thus each trial started with 30 sec isolated action of the tone CS giving 250 'test trials' during which any imagery reported might be inferred to be due to the conditioned properties of the tone; 242 of these trials were continued as training trials being reinforced by the presentation of the illuminated circle and cross target. Eight trials were not reinforced. On these trials the tone was terminated by the experimenter after 60 sec unless the subject was reporting visual imagery, when the tone was allowed to continue until terminated by the subject in the normal way. These trials were inserted in the schedule in place of the 10th, 50th, 80th, 120th, 200th and 250th trials.

Subsidiary experiments not reported here entailed a further 18 trials of the normal training type being employed over days 26–29. Thus each subject received a total of 260 'test'/training trials with a prolonged CS during which conditioned responses with less than 30 sec latency could appear, and 8 unreinforced (tone only) test trials in the course of 29 experimental sessions.

The subjects were not aware that the experiment was concerned with conditioning. They were

Table 1. *Number of images experienced per day in response to tone and before the flash i.e. conditioned afterimages**

Day	Subjects 1	2	3	4	5	6	Total
1	—	—	—	—	—	—	—
2	—	2	—	—	—	—	2
3	—	1	—	—	—	—	1
4	—	—	—	—	—	6	6
5	—	7	—	—	—	11	18
6	—	4	—	—	—	10	14
7	—	8	—	—	—	11	19
8	—	8	—	—	—	10	18
9	6	10	—	—	—	11	27
10	7	13	3	—	—	10	33
11	8	14	4	—	—	9	35
12	14	20	3	—	—	13	50
13	12	16	2	—	—	8	38
14	6	15	3	—	—	10	34
15	13	16	4	—	—	28	61
16	10	11	4	—	—	10	35
17	6	16	1	—	—	10	33
18	11	1†	5	—	—	9	26
19	7	3	3	—	—	10	23
20	10	5	1	—	—	25	41
21	11	7	—	—	—	10	28
22	7	8	4	—	—	10	29
23	9	13	5	—	—	10	37
24	8	1	4	—	—	10	23
25	10	10	4	—	—	17	41
Total	155	209	50	0	0	258	672

* Number of images exceeds number of trials per day as some subjects experienced more than one image during the 30 sec of tone.

† Resumption after 11 weeks of illness.

Table 2. *Number of conditioned images reported on unreinforced test trials*

Subject	Trial number 10	50	80	120	150	200	230	250
1	—	—	—	2	1	3	3	3
2	—	2	3	6	5	2	—	—
3	—	—	—	—	—	1	1	—
4	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—
6	—	3	1	6	19	15	7	7

told that it was an investigation of the effects of tone on the visibility of afterimages, and they were in no way led to expect to see images in response to tone alone before they had reported such images three times. Once they had reported images during the tone and in advance of the flash on 3 trials they were told that 'This result has been mentioned to me before. It is very interesting and I would like to pay you particular attention to it should it occur again.' Reasons for including this facilitatory instruction will be discussed.

No control group was employed in this experiment as adequate controls had been run previously in similar experiments (Davies, 1974*a*, *b*). In these experiments two groups of five subjects had been maintained on a Rescorla type schedule (Rescorla, 1967) with ten 'pairings' a day for 20 days: one group with a 2·5 sec tone, the other with a 25 sec duration tone. Not one subject in either of these groups reported imagery in response to the tone alone on any trial.

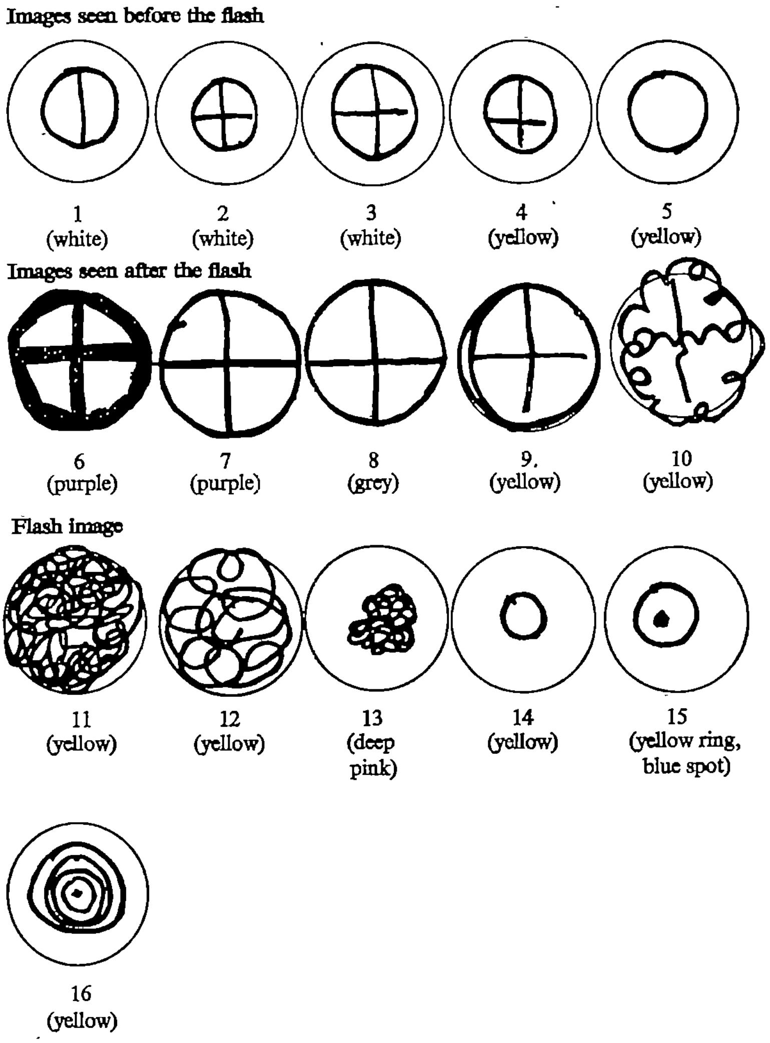


Fig. 1. Drawings made by Subject 6 following trial 148. The faint outline circle indicates the 'normal' size of the target image. The numbers give the sequential order, the colour refers to the image not the ground upon which it appeared.

Results

Four of the six subjects experienced images in response to the tone in advance of the presentation of the illuminated target on some of the 242 combined 'test'/training trials; the same four subjects experienced images in response to the tone on some of the eight trials not reinforced by flash. These results are summarized in Table 1 and Table 2 respectively. A fifth subject experienced images in response to the tone and before the flash on 12 of the 18 normal 'test'/training trials presented during the 4 days subsequent to the main experiment. A selection of the images experienced by one subject is presented in Fig. 1.

The *number* of conditioned images is relatively uninformative however. It indicates the presence but not the strength of conditioning as the first such images to appear

Table 3. *Mean latency and duration of conditioned images**

(First 5 days on which such responses were reported. L = mean latency (sec); D = mean duration (sec).)

	Day 1		Day 2		Day 3		Day 4		Day 5	
Subject	L	D	L	D	L	D	L	D	L	D
1	15·2	8·9	18·5	9·0	15·8	9·5	12·6	13·6	17·0	9·7
2	9·3	7·3	22·8	2·5	19·7	2·9	18·5	2·4	17·3	4·4
3	17·0	13·0	4·3	25·7	16·5	13·5	17·0	13·0	13·6	16·2
4†	15·0	15·0	26·8	3·2	0	0	14·4	15·6	—	—
5	—	—	—	—	—	—	—	—	—	—
6	16·5	11·5	14·4	17·2	12·8	17·2	11·5	18·5	11·0	18·9

* Excluding conditioned images occurring on tone only test trials.
† Reports for 4 days only (see text).

are characteristically labile and of short duration; although not invariable the subject's first reports are usually of a number of ill-defined images that change, or come and go, rapidly. The latency and duration of the conditioned images are better quantitative indicators of strength. Table 3 shows the mean latency and mean duration of the conditioned responses for the first 5 days on which these appeared for each subject. Only one latency is recorded on each trial, that from the onset of the tone CS to the beginning of the first conditioned image; the mean duration was obtained by summing the duration of all conditioned images and dividing by the number of trials on which these occurred. Subjects 1, 3 and 6 show an increase in mean duration over the first 5 days during which conditioned images were experienced, subject 2 shows a similar increase over the period day 2–5. For these four subjects this trend is significant at the 0·05 level (Page's L test; $L = 197, P < 0{\cdot}05$). Over the period covered by Table 3 the latency shows no systematic change; however, over the whole experiment the general trend was for the latency to decrease (for images to be reported nearer the tone onset) and for the duration time for all images, both those reported in advance of the flash and those reported after it, to increase. The duration data for the conditioned images occurring during the 30 sec before the flash are misleading as conditioned imagery may only be identified unequivocally as such during this period. After the target has been illuminated it seems reasonable that unconditioned images should occur. The subjects' reports confirm this as they inevitably reported a marked colour change in the image as the target was flashed (see Fig. 1). However, the imagery presumably continues as a mixture of unconditioned and conditioned imagery; only subjects who later conditioned showed increased duration of imagery after the flash. This portion of the response shows an increase before any response to tone alone is reported. Once anticipatory images, i.e. conditioned images, were reported they lengthened rapidly not only in the direction of the tone onset but also toward the flash presentation. It is for this reason that Table 3 is restricted to 5 days; during this period most of the images had terminated before the flash was presented whereas later they had not.

The evolution of the conditioned images is seen more clearly from the duration of imagery on the trials which were not reinforced, i.e. 10, 50, 80 *et seq.*, where more than 30 sec of purely conditioned imagery could be reported. These data are reported in

Table 4. *Duration of conditioned imagery on tone only test trials*

(Duration (sec) measured from outset of first image reported to off-set of last image reported.)

Subject	Trial number							
	10	50	80	120	150	200	230	250
1	—	—	—	54	44	94	104	143
2	—	117	142	115	86	54	—	77
3	—	—	—	—	—	85	67	—
4	—	—	—	—	—	—	—	—
5	—	—	—	—	—	—	—	—
6	—	168	52	260	423	190*	150*	150*

* Tone stopped by the experimenter as this subject reported imagery as long as the tone sounded up to 10 min 23 sec on trial 160. The images terminated within 3 sec of the tone off-set at 10 min 20 sec.

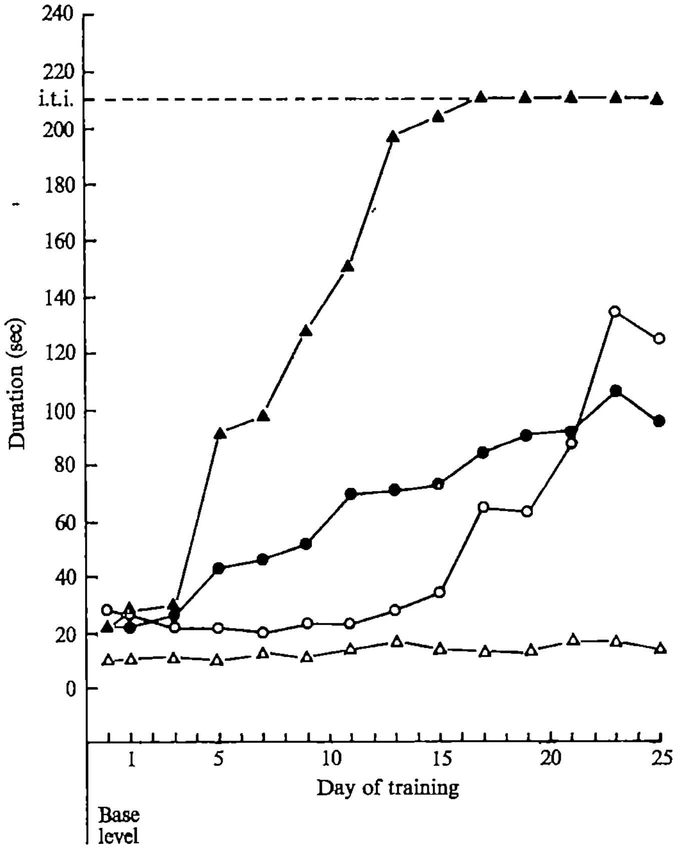


Fig. 2. Mean duration of afterimagery: alternate days; group data, subject 1, 5 and 6. Data for Subject 2 omitted – protracted break in conditioning schedule affected the afterimage durations. ●—●, Group (*n* = 5); ○—○, Subject 1; △—△, Subject 5; ▲—▲, Subject 6.

Table 4. Subjects 1 and 6 show the most marked increases; the full extent of Subject 6's increase is not shown as it became necessary to terminate the trials by turning the tone off. This subject came to experience imagery in response to the tone for as long as the tone sounded up to 10 min and 23 sec (tone stopped by experimenter at 10 min 20 sec on trial 160).

The mean daily duration data indicate a marked increase in the duration of all imagery, though this effect is not so marked as that reported by Bzhalava (1958, 1965). By day 25 the mean duration of imagery on all trials (post-flash imagery only) increased fourfold for the group as a whole (base level for unconditioned afterimagery, 20·0 sec; mean duration on day 25, 84·4 sec, $n = 6$ in both cases). That this increase is associated with the evolution of conditioning may be inferred from the fact that Subject 5, who did not condition, maintained a stable mean daily duration throughout the experiment; Subject 4, who did not condition until day 26, showed very slight increase. This may be seen from Fig. 2.

DISCUSSION

The subjects were given facilitatory instructions, designed to encourage the reporting of any visual effects, *after they had spontaneously reported images* in response *to the tone alone* on three trials. Previous, unreported experiments had indicated that subjects clearly associate visual effects with visual stimuli and are unwilling to report apparently irrelevant visual effects. One subject in an early experiment reported that the flash had failed again, did I want him to report what he was seeing? Post-experimental questioning revealed that he had experienced quite clear images in response to tone alone on at least 20 trials before he thought the matter worth mentioning. The present instructions do not avoid this danger, that the subject will not report visual effects in the absence of a prior visual stimulus, but they do minimize the chances that the subject will stop reporting them once he has started to do so. Unless such instructions are given at the outset it seems inevitable that some data must be lost, but this seems preferable to the possibility that subjects may respond to experimenter expectations and 'manufacture' responses.

The superiority of this method for conditioning afterimages over that employed previously is demonstrated by the results. Five of the six subjects conditioned, two of them strongly. Once conditioning was well established the descriptions of the conditioned images were virtually indistinguishable from descriptions of unconditioned images (see Fig. 1; images seen before the target was illuminated are more clearly defined than many images seen after the flash). Subjects 1 and 3 required many more trials than Subjects 2 and 6, but all were strongly conditioned. The least strongly conditioned subject was Subject 4, who described her most clearly defined conditioned image as, 'Light coloured circle, nothing in the centre. Parts are speckled black, these moving in to the centre.' This compares well with the most clearly defined conditioned images reported previously.

The earliest conditioned images reported were merely vague or circular patches of light the same size as the normal images of the target. These descriptions seem to be related to Bzhalava's second level of evolution of conditioned images described as 'relatively stable sensations of an amorphously illuminated grey field' (Razran, 1971). Possibly the reports correspond with his third level, 'square outlines of the field (perception)' (Razran, 1971). Subjects 1, 2, 3 and 6 may have reached the highest of the five levels of evolution reported by Bzhalava, i.e. 'clear perception of...the US' as they reported clear images of a positive circle and cross shape, white or yellow, against a dark or black background in response to the tone alone. None of the subjects reported experiences comparable to the lowest level described by Bzhalava (1965): 'fleeting amorphous light sensations "resembling moving clouds"' (*op. cit.*), which suggests that the identification of the earliest stages of a conditioned image may depend upon the use of trained observers or facilitating instructions given at the outset.

Of the six subjects only Subject 6 showed an increase in duration of afterimagery equal to those reported by Bzhalava: by the 160th trial this subject's image durations were strongly bound to the tone and would continue for as long as the tone was sounded (up to 10 min 20 sec, the longest duration tested). The response was also remarkably stable: when a subject in a psychophysical experiment, for a different experimenter, this subject reported visual imagery in response to a 1000 Hz tone some 10 months after the original conditioning experiment. The form of imagery reported on this occasion resembled that previously reported on days 14 and 15 of this experiment (see Fig. 1).

The subject who did not condition, Subject 5, may have failed to do so for any of several reasons. Unknown to the experimenter (until the 20th day of training) she was in the early months of pregnancy, taking a prescribed sedative to aid sleeping, an anti-nauseant to combat morning sickness, and was scheduled for morning conditioning sessions. Costello (1963) reported that depressant drugs shorten unconditioned afterimages: Malitz & Kanzler (1970) confirm that tranquillizers and depressants decrease sensitivity to afterimages; and the Popovs (1953*a*, *b*) reported that alcohol (a depressant) inhibits both normal and conditioned afterimages.

CONCLUSIONS

The results indicate that prolonging the tone so that it is paired not only with the unconditioned stimulus but also with the unconditioned response is a very effective configuration to bring about this form of conditioning. The subjects were more strongly conditioned using this paradigm than in any reported using brief auditory stimuli. However, none were as strongly conditioned as those subjects reported by Bzhalava if we utilize only the increasing duration of afterimagery as an indirect measure of conditioned response strength. The reasons for this discrepancy are a matter for further investigation. Nevertheless the data suggest that conditioned imagery may be more readily obtainable when the extinction effect of normal test trials is minimized by combining the test and training trials.

This research was supported by a Departmental Studentship of the Psychology Department of the University of Aberdeen. I am also indebted to Dr N. E. Wetherick for his comments and advice.

REFERENCES

BZHALAVA, I. T. (1958). Natural-science foundations of psychological action. Tbilisi: Metsniyerba. (Summarized in Razran, 1971.)

BZHALAVA, I. T. (1965). *Perception and Set.* Tbilisi: Metsniyerba. (Cited in Razran, 1971.)

COSTELLO, C. G. (1963). The effects of meprobamate on the visual afterimage. In H. J. Eysenck (ed.), *Experiments with Drugs*. Oxford, London: Pergamon.

DAVIES, P. (1974*a*). Conditioned visual afterimages: I. *Br. J. Psychol.* **65**, 191–204.

DAVIES, P. (1974*b*). Conditioned visual afterimages: II. *Br. J. Psychol.* **65**, 377–393.

MALITZ, S. E. & KANZLER, M. (1970). Effect of drugs on perception in man. In Hambourg, Probram & Stunkard (eds), *Perception and its Disorders*. Baltimore: Williams & Wilkins.

POPOV, N. A. & POPOV, C. (1953*a*). Contribution à l'étude des fonctions corticales chez l'Homme; par la méthode des réflexes conditionnées électrocorticaux. Action de l'alcool sur les images consécutives et leur conditionnement: I. *C. r. hebd. Séanc. Acad. Sci., Paris* **237**, 930–932.

POPOV, N. A. & POPOV, C. (1953*b*). Contribution à l'étude des fonctions corticales chez l'homme par la méthode des réflexes conditionnées électrocorticaux. II. De la modification per l'alcool des couleurs des images consécutives et des images consécutives conditionnées. *C. r. hebd. Séanc. Acad. Sci., Paris* **237**, 1439–1441.

RESCORLA, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychol. Rev.* **74** (1), 71–80.

RAZRAN, G. (1971). *Mind in Evolution*. Boston: Houghton Mifflin.

(*Manuscript received* 14 *June* 1974; *revised manuscript received* 23 *May* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 191–201
Printed in Great Britain

THE ROLE OF INFORMATION-PROCESSING CAPACITY IN THE CONSERVING OPERATION

VERNON HAMILTON AND GILLES LAUNAY
Reading University

Previous studies had shown that success in solving conservation of quantity problems varies with the amount of information contained in the problem, and that there are individual differences in coping with increasing informational loads. Amount of information was defined by the number of equational steps required for the solution of a problem. In testing the validity of this definition, it was predicted that criterion groups of subjects would differ significantly in other test situations measuring information-processing capacity. Groups of 10 non-conservers, 30 poor conservers, and 47 good conservers were selected from a subject pool of 150 aged between 6 and 9 years. The validation was carried out on card-sorting tasks differing in binary choice demands, and on coding tasks differing in the number of primary stimuli and the number of recoding steps. Most of the predictions received highly reliable confirmation.

The studies to be reported here form part of a long-term programme of work on the general relationship between intrapersonal stress and the development of effective cognitive processes. They arose in the first instance from findings on the impairment of the basic cognitive skills of size constancy and conservation in groups of schizophrenics (Hamilton, 1963, 1966, 1972*a*), and from more recent findings that the measurement of fine-grain differences in conservation reliably distinguishes between groups of children whose mothers were assessed as respectively 'accepting' or 'rejecting' (Hamilton, 1971, 1972*c*; Hamilton & Moss, 1974). Mothers, as well as children, in the latter experiments were 'normal' by all the usual social and psychological criteria. The link between the two sets of studies is provided by a general statement which says that attention to, and the processing of, information that is relevant to the demand characteristics of a given, objective situation may be suboptimal in people who have experienced socialization stress at important stages in their early lives. The point at which this statement can be directly tested has not been reached yet, because we are at the intermediate stage of testing the validity of the assumption that deficits in basic cognitive skills, such as conservation, are substantially due to a general suboptimal capacity for processing external stimuli.

It has been demonstrated that conservation of quantity may be redefined as a continuous variable for which individual differences can be ordinally described (Hamilton, 1972*b*; Hamilton & Moss, 1974), and that groups of subjects assumed to differ in socialization stress and anxiety perform with different levels of competence on a conservation task (Hamilton, 1972*c*; Hamilton & Moss, 1974). From the outset it was necessary to refine our measure of conservation so that its information-processing requirements at least had face validity, and only then attempt a concurrent validation of the information-processing assumption by reference to performance on other tasks with which information-processing terminology is more orthodoxly associated. The conserving capacity had been selected for special attention because it develops at the time when socialization pressures on children are normally increasing, and because adequate development of the concept of invariance is probably one of the most fundamental logical skills in adult reasoning.

The re-analysis of a series of conservation of quantity problems of systematically increasing difficulty in terms of information content has been completed. Details concerning the method are given elsewhere (Hamilton & Moss, 1974), so that a summary must suffice here. A first attempt to demonstrate a wide range of individual differences in conservation, as well as the continuity of this cognitive skill, yielded a negatively accelerated subject pass rate curve (Hamilton, 1971, 1972*b*). Problem difficulty was manipulated by gradually increasing the number of originally equal amounts (standards), the number of distributions and redistributions of given quantities, and the number of separate parts formed by this procedure. At this stage, only an *ad hoc* method was available for the quantification of within- and between-problem levels of difficulty. Inspection of the relationship between performance and complexity of the display parameters suggested that problem difficulty might depend on the number of steps necessary to achieve an invariance or equality response. The steps were defined as equations embedded in, and presented by, the distribution characteristics of a problem. The greater the number of originally equal quantities, the number of distributions and redistributions, and the number of separate parts formed thereby, the greater would be the number of steps that would have to be processed on the way to a conserving response. The greater the number of steps presented by a problem – and thus its information load – the higher, it was assumed, the difficulty level of the problem. An additive index of problem difficulty was derived which is a function of the above parameters, the number of necessary reversibility operations, and the number of undistributed but simply deformed materials. Our formula permitted the design of conservation problems which operationally differed in difficulty, and where between-problem intervals were quantified. It was argued that a so-called immediate awareness that 'nothing was added or taken away', was neither 'immediate' nor an invalidation of the procedure. This kind of justification of a conservation response could be made only after a rapid search–scan–and–match sequence of steps had been carried out. The conceptual linking of equational steps – or *chunking* of information – by some subjects probably would be reflected in faster reaction times, and would only indicate the presence of higher order and more mature analysing processes. The range of the calculated steps in our series was 3–68, with between-problem intervals of 2–7 steps. The correlation of percentage pass rate and problem difficulty, as defined above, was $-0{\cdot}941$. We propose that this result confers face validity upon the equational variable in the conservation test series.

The present report is concerned with supplying concurrent validity. The dependent information-processing variables chosen were (1) speed of sorting playing cards into categories of decreasing degrees of redundancy, and (2) errors and speed in a coding task with substitution requirements of increasing length. All three tasks demand a stepwise and additive analysis of stimuli, with decision and choice points pathways of increasing length which could be predicted to make increasing demands on an information-processing system (Newell & Simon 1972). For the sake of simplicity of presentation, the card-sorting and coding experiments will be reported in separate sections.

EXPT. I: CONSERVATION AND CHOICE TIME WITH VARIATIONS IN REDUNDANCY

Method

Information theorists have demonstrated on adult subjects that the general processing capacity of the human operator is finite (Miller, 1956), so that errors in performance will increase as the limit is approached, and that speed of handling information is a function of redundancy. As the number of binary choices in a task increases, so does the decision time (Crossman, 1953). Little research has been done on children, but some evidence is available to show that the capacity to handle redundancy is a function of age and level of conceptual development (e.g. Osler & Kofsky, 1965). It is not suggested that there is an identity between the type of information-processing assumed to enter into solving conservation problems and that which determines binary choice response time in situations requiring the transmission of an increasing number of 'bits'. Where task complexity is defined by number of binary choices, however, speed shows a monotonic relationship to complexity. A similar relationship between speed and complexity applies to other tasks. It is possible that levels of conceptual operations are substantially affected by processing speed, since short-term memory decay will increase with slowness, and so limit the number of stimuli or concepts on which central processes can operate. The assessment of choice response time is thus an assessment of an important component of total information-processing capacity. The following hypotheses were tested:

(1) High scoring conservers compared with non-conservers and poor conservers are faster in card-sorting tasks containing varying amounts of redundancy.

(2) The superior card-sorting speed of high scoring conservers increases with decreasing amounts of redundancy.

Material

A playing-card sorting task employed by Crossman (1953) was used in this experiment because the task is simple, the test material familiar, and predictions of performance could be anchored in available data. Crossman used the following five sorting tasks (with the number of 'bits' – $\log_2 n$ – in parentheses): Red/Black from a prearranged pack (0), Red/Black from a well-shuffled pack (1), Four Suits (2), Six and Below in Suits, and Seven and Above in Suits (3), Ace to King in Order, Red/Black (4·70). A pilot study yielded highly significant differences between competent and less competent conservers for these sorting categories. Unexpectedly, however, condition $\log_2 n = 3$ led to abnormally high sorting times which were not found for the adult subjects of Crossman's original study. Re-examination of the definitions of the sorting categories for this condition suggested that they were more complex than the definitions of other sorting categories, and that the binary choices were not immediately obvious for children because dual criteria had to be applied. The task requirements were thus different from those reflected in information-transmission concepts, requiring additional processes ('*what* is below six...'*what* is above seven'), so that the playing card sorting categories required manipulation in order to yield a monotonic redundancy × sorting time relationship.

A subsidiary experiment was required, therefore, to discover a set of sorting categories with homogeneous choice demands. For this purpose a parallel set of sorting tasks was devised embodying for each level of redundancy dual criterion sorting categories. The original ('Standard'), and new ('Complex') sorting tasks were arranged in alternating order so that each subject performed eight separate playing-card sorts, preceded by sorting a Black/Red alternating pack ($\log_2 n = 0$). The sorting categories of the 'Complex' series as graphically presented to subjects on marker cards were: 1, 2, 3, 4, 5, 6 and 7, 8, 9, 10, J, Q, K (1·015); red 1, 2, 3, 4, 5, 6 – black 1, 2, 3, 4, 5, 6 – red 7, 8, 9, 10, J, Q, K – black 7, 8, 9, 10, J, Q, K (1·985); Diamonds/Hearts 1, 2, 3 – Clubs/Spades 1, 2, 3 – Diamonds/Hearts 4, 5, 6...Clubs/Spades 10, J, Q, K...(3·000); Clubs 1, K – Clubs 2, Q – ...Hearts 3, J...Clubs, 9, 10 – repeated for all Suits, with the addition of Clubs and Spades 7 – and Hearts and Diamonds 7 – (4·836). The Suit symbols, colours and denominations were neatly drawn on category marker cards.

Table 1. *Subject parameters*

				Educational quotient*†	
	n	Age	Conservation score	Reading	Arithmetic
Non-conservers	10	$\bar{x}$ 7–2	$\bar{x}$ 0	$\bar{x}$ 117·90 (6)	$\bar{x}$ 109·81 (4)
		S.D. 1–0	S.D. 0	S.D. 10·27	S.D. 6·76
		Range 6–3 to 9–3	Range 0	Range 104–130	Range 101–120
Low scoring conservers‡	30	$\bar{x}$ 8–2	$\bar{x}$ 3·13	$\bar{x}$ 112·33 (27)	$\bar{x}$ 103·74 (25)
		S.D. 0–81	S.D. 1·76	S.D. 11·35	S.D. 8·97
		Range 7–0 to 9–5	Range 1–7	Range 95–140	Range 92–121
High scoring conservers¶	47	$\bar{x}$ 8–2	$\bar{x}$ 15·23	$\bar{x}$ 111·21 (43)	$\bar{x}$ 105·92 (35)
		S.D. 0–95	S.D. 2·11	S.D. 15·51	S.D. 10·73
		Range 6–1 to 9–9	Range 9–17	Range 85–150	Range 91–129

* Primary Reading Test 2 and Non-Verbal Test 5 – both NFER.
† In parentheses, number of subjects yielding means.
‡ Conservation scores of 7 or less.
¶ Conservation scores of 8 or more.

Procedure

The category markers were drawn on rectangular cards and were placed on a table in a regular display prior to sorting into categories. All sorting was done with cards facing upwards, all packs of playing cards were new, and well shuffled except for $\log_2 n = 0$. Two practice sorts into four and six unspecified groups preceded the experimental conditions with cards facing downwards from which movement times per card sorted were calculated. These times were subtracted from the sorting times per card to obtain processing or decision times. Where necessary, error corrected sorting times were obtained by adding for each error to actual sorting time twice the sorting time per card obtained in practice sorting. In order to control for search time, the number of symbols identifying sorting categories on the marker cards, and the visual angle subtended by the marker card display on the experimental table, were held as constant as possible within and between the 'Standard' and 'Complex' sorts. Subjects were instructed to sort the cards into the appropriate categories without making any errors as quickly and carefully as possible, and specifically not to correct any errors they might make.

Subjects

These were selected from a pool of 150 subjects whose conserving ability had been assessed previously (Hamilton & Moss, 1974). A tripartite division of conservation scores, obtained from children who remained available for further experimentation, yielded 47 good conservers and 40 non-conservers and poor conservers. Intelligence was not formally assessed, but measures of educational capacity were obtained. These data are incomplete because of protracted illness of some children. None of the differences between educational quotients was significant.

For the subsidiary experiment testing the homogeneity of processing demands in sorting tasks decreasing in redundancy, subjects were taken as a single group of 87. For the validation studies, subjects were divided into three criterion groups with non-overlapping scores on our graphical test of conservation of quantity which was described above. The basic group parameters are summarized in Table 1. While the direction of our studies towards the measurement of fine-grain individual differences of cognitive processes suggested a greater number of criterion groups, this procedure was made inappropriate by the bunching of scores at the easy (scores of 7 or less) and difficult (scores of 13 or more) ends of the conservation test series.

Results

Fig. 1 describes the findings of our subsidiary experiment, with the addition of Crossman's (1953) original findings. They confirm that Crossman's category descriptions ('Standard') are non-monotonic for children, and that our own series of categories ('complex') yields a virtually linear decision time × bits function.

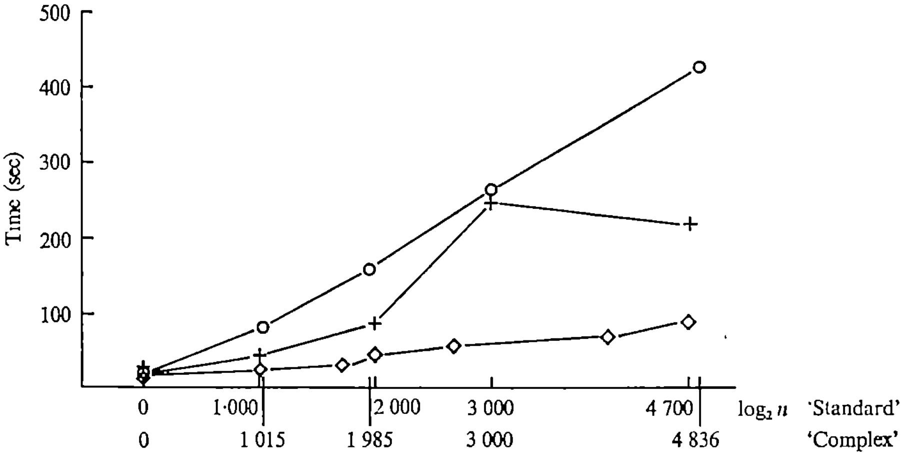


Fig. 1. Mean decision times (52 cards) 'Standard' and 'Complex' sorting categories. +—+, 'Standard'; ○—○, 'Complex'; ◇—◇, Crossman's (1953) data (estimated).

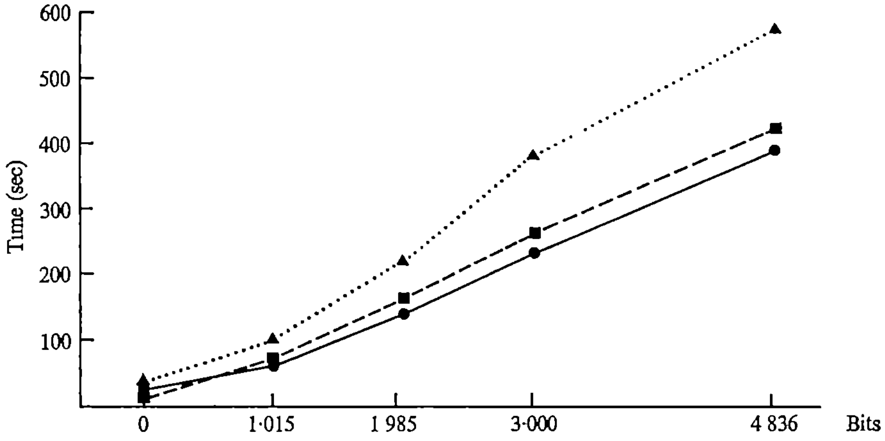


Fig. 2. Speed of information-processing and conserving capacity. ●—●, Good conserver ($n = 47$); ■—■, poor conserver ($n = 30$); ▲··▲, non-conserver ($n = 10$).

With this evidence we were able to test for the predicted relationship between conserving capacity and binary choice decision time for four conditions of redundancy. The results for good, poor, and non-conservers are shown in Fig. 2. They clearly support our first two hypotheses in terms of the main effect ($F = 6{\cdot}33$, d.f. = 2, 84, $P < 0{\cdot}01$), and in terms of a significant groups × conditions interaction ($F = 4{\cdot}03$, d.f. = 8, 336, $P < 0{\cdot}0001$).

All t tests between non-conservers and poor conservers, and between non-conservers and good conservers are significant. A separate analysis of variance for the data from poor and good conservers indicated non-significant main and interaction effects, even though the separation between the mean decision times of the two groups showed a progressively increasing degree of difference. Only one t test for 1·985 bits was significant at the 5 per cent level. The demonstrated main effect may be influenced substantially, therefore, by the age factor, and by differences in the level of concept acquisition of the playing-card categories of the younger group. Of all the possible correlations with $n = 87$, 56 per cent were moderately significant, the highest value being 0·328 between number of conservation problems solved correctly and 1·985 bits.

EXPT. II. CONSERVATION AND CODING ABILITY WITH VARIATIONS IN NUMBER OF RECODING STEPS

Method

In the present context, the term coding is defined, in traditional terms, as the activity or ability to substitute one particular sign or symbol for another before selecting an appropriate response. It is possible to redescribe symbol substitution in terms of an equation, since its functional significance is an identity statement. Thus, the traffic light symbol 'Green' equals 'Go', and in the coding subtest of the *WISC* 9 equals ⊣. In each case the symbol must be cognitively identified and matched with a functionally equivalent associate. We are making the assumption that this activity could serve to define the information-processing demands of coding tasks, if the amount of coding is defined by the number of equational steps presented in a task display and by the experimental instructions. Systematic manipulation of the number of equational steps, and thus task difficulty in terms of information-processing, can be achieved by varying (1) the number of stimuli presented simultaneously for coding operations, and (2) the number of times each of these stimuli requires recoding before an appropriate response can be selected. We, therefore, designed an experimental display containing these requirements. The following hypotheses were tested:

(1) High scoring conservers compared with non-conservers and poor conservers make fewer errors and are faster in coding tasks.

(2) The superiority in coding capacity of high scoring conservers increases as the substitutional demands of the coding task increase.

Material

The major piece of apparatus was a rectangular box 42·0 cm long, divided into six cubicles measuring 4·5 × 4·5 cm. Each cubicle could be illuminated separately by a 24 V, 2·8 W bulb. On the ground-glass screen fronting the box, six squares outlined by black tape, and the upper-case letters A–E 4·0 cm high were marked out. Each letter could be illuminated singly or in combination with any of the others by feeding into a photoelectric tape reader (Lectromec MK 4) a 22 opaque channel film strip on which the required sequence of primary stimuli was punched out. Activation of the tape reader motor presented a single sequence of signals to subjects, simultaneously starting an electronic timer (Advance SC3). After each response, activation of the reader motor extinguished the last set of primary stimuli, switched to the programme for the next stimuli, and reset the timer. Arranged in rows below the letter display, and in random sequences, were the rectangular cards giving the codes linked to the letters. The simplest task – one recoding stage – required the substitution of colours: yellow, red, green, for letters. Also arranged in rows were 'rule' cards indicating that: 'yellow is more than red', 'red is more than green', 'yellow is more than green', 'two yellows are more than two reds', 'two reds are more than two greens', 'two yellows are more then two greens', and 'two reds are more than one yellow'. Subjects had two response keys marked 'yellow' and 'red', and were required to press the key appropriate to the colour code 'values' of the illuminated letters. Thus, if the three letters illuminated required substitution into 'yellow', 'red', 'red', pressing the 'red' response key signified the correct response. In order to increase the task difficulty, two further recoding stages were designed:

letters → geometrical symbols → colours, and letters → geometrical symbols → numerals → colours.

Fig. 3 describes the task for two recoding stages. All subjects proceeded from one to two to three recoding stages, responding to three lists of 21 primary stimuli (letters) which counterbalanced the number of times 'yellow' or 'red' signified the correct response. This was achieved by introducing the code of 'green' which always appeared in combination. Each list of primary

Primary stimuli	Upper case Letters A–F
	↓
First recoding stage	Geometric symbols – in random order: B and D = rectangles, A and E = circles C and F = stars (6 coding key cards)
	↓
Second recoding stage	Colours – in random order: Circles = yellow, Rectangles = green Stars = red (6 coding key cards)
	↓
Rule application	Yellow > red > green, etc.
	↓
Response key	Yellow or red

	SPECIMEN PROBLEM
Primary stimuli	illuminate letters C D E F
	↓
Recode	STAR RECTANGLE CIRCLE STAR
	↓
Recode	RED GREEN YELLOW RED
	↓
Rule	2 REDS 1 GREEN 1 YELLOW 2 REDS + 1 GREEN > 1 YELLOW
	↓
Response	RED

Fig. 3. Coding task display for two recoding stages: potential display characteristics.

stimuli contained 2 single letters, 5 groups of two letters, 6 groups of three, 5 of four, 2 of five, and 1 of six letters, and a different list was used for each of the three recoding stages.

The simultaneous variation of the number of primary stimuli (letters) and recoding stages (substitutions) yields an ordinal scale for the number of steps or equations that need to precede response selection. For example, our specimen problem in Fig. 3 requires the processing of three equations for each of the four primary stimuli. We arrive at this as follows: C = Star, Star = Red, and the deduction that C = Red. Thus, four simultaneously presented primary stimuli required 4 × 3 equational operations. With only a single recoding stage, the number of steps is equal to the number of primary stimuli. With three recoding stages we need three substitutions and two deductions, e.g. C = Star, Star = 4 (say), 4 = Red, Star = Red and C = Red. A table of operational or equational steps can thus be constructed for any combination of primary stimuli and recoding stages from which coding problems systematically increasing in processing requirements can be selected (Table 2). This analysis gives us a definition of difficulty of a coding task in information-processing terms which is analogous to that which had been applied to conservation tasks (Hamilton & Moss, 1974): the larger the number of primary stimuli and recoding stages, the greater the number of operational or equational steps. The greater the number of steps, the greater the information-processing demands.

The experimental design was necessarily constrained by an age range of more than three years covering the youngest and the oldest subjects, by the need to retain subjects' involvement and cooperation by a task of reasonable length, and by avoiding response bias through place learning with two response keys. Preliminary testing suggested that a primary stimulus series of six signals, and three recoding stages would yield meaningful performance ranges.

Table 2

Primary stimuli Recoding stages	1	2	3	4	5	6
1	1	2	3	4	5	6
2	3	6	9	12	15	18
3	5	10	15	20	25	30

Procedure

The coding tasks were laid out on the experimental table directly in front of the primary stimulus generator. They were black ink drawings on white cards, presenting in the simplest possible form the substitutions for each primary stimulus, i.e. A is ○; ○ is RED; or B is □, □ is 3, 3 is YELLOW, etc. For each series of recoding stages, the codes for all primary stimuli were available, even if not all stimuli were activated. The colour 'value rules' were also clearly drawn on white card, and gave sufficient information to deduce a correct solution. We did not feel able to fit a quantifying method to the transitivity operations involved in rule application, and would offer the argument that individual differences in an established transitivity skill will emerge only if the information to which this operation has to be applied comes close to subjects' upper information-processing limits defined here by the number of preceding equational operations. There is good evidence that differences in understanding the rules of the task did not determine the direction of our findings, since errors made by all three criterion groups were relatively few with low levels of information-processing. The full signal display remained available to subjects until they had responded. While the absence of a time limit was intended to deal with the gross effects of short-term memory loss, it is not unlikely that slow processers are at a disadvantage, even though memory decay may be reduced by re-examination of the display. On the other hand, we expected that fast responses would often occur in conjunction with errors, and, therefore, laid greater weight in our predictions on errors than on reaction times – the two dependent variables. The independent variable was scores on our test of conservation.

Subjects

These were the same groups of 47 good conservers, 30 poor conservers and 10 non-conservers who took part in the card-sorting experiment and whose basic characteristics are described in Table 1. The instructions to subjects were to change the letters (primary stimuli) into colours (one recoding stage), and to 'work out which of the buttons – yellow or red – you will have to press using the rules written on the cards'. Two examples were worked out with the children for whom adequate reading ability had been ascertained at an earlier stage. Very few difficulties of comprehension were observed. The instructions were modified appropriately when more than one recoding stage was presented.

Results

The most descriptive findings relating conserving capacity to coding with an increasing number of primary stimuli and recoding stages for the three experiment groups, are illustrated in Fig. 4. As predicted, errors increased with the processing demands of the tasks. For the analysis of the effects of the number of primary stimuli we have confined ourselves to three categories (2, 3 and 4 stimuli) because of in-

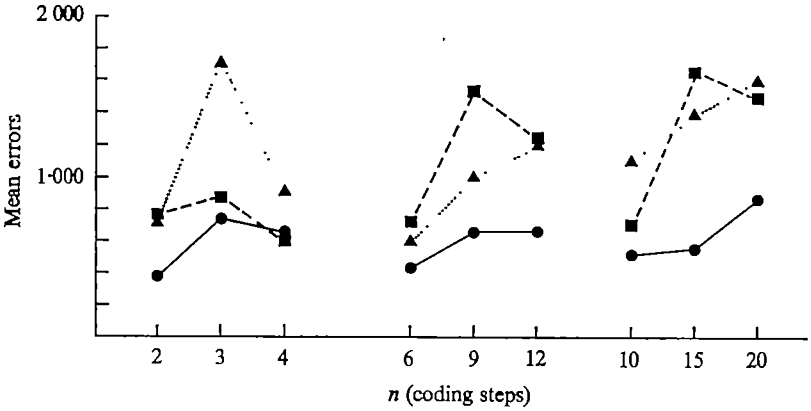


Fig. 4. Recoding and conserving capacity. ●—●, Good conservers ($n = 47$); ■—■, poor conservers ($n = 30$); ▲··▲, non-conservers ($n = 10$).

sufficient data from the other categories which, as described above, were less frequently presented. Analyses of variance showed significant groups and conditions effects ($F = 9{\cdot}52$, d.f. $= 2, 84$; and $F = 6{\cdot}25$ d.f. $= 8, 672$, $P < 0{\cdot}001$), and a significant groups × conditions interaction ($F = 2{\cdot}16$, d.f. $= 16, 672$, $P < 0{\cdot}01$). Very similar levels of significance were found when analysis of variance was restricted to poor and good conservers. Reaction time differences failed to reach significance, but were in the predicted direction: the greater the processing demands of the task, the slower the response. The *t* test analyses indicated, as already apparent from Fig. 4, that the main and the interaction effects for errors are not largely due to the expected lower coding competence of the significantly younger non-conservers. A larger number of significant and higher *t* values differentiated the good and poor conservers, than the former and the non-conservers. Only a single significant *t* value was obtained for the difference between poor and non-conservers. Of all the possible correlations for $n = 87$, 50 per cent were moderately significant, the highest value being 0·380 between number of conservation problems solved correctly and the sum of coding errors for problems with 6, 9 and 12 substitutional steps.

Discussion

Our hypotheses predicted superior performance for good conservers compared with poor and non-conservers. Non-conservation on our test was regarded as a zero score, and in a preliminary analysis good conservers were compared with non- and poor conservers combined. This comparison yielded highly significant main and interaction effects for both experiments. Because the average age of non-conservers emerged as significantly lower, the data presented here give an account of the performances of three rather than two groups. While age clearly contributed to inferior card-sorting speeds of the non-conservers, Table 1 indicates that the overlap in the age range of all three groups was considerable, while educational quotients, and thus probably intelligence, were comparable. A more clearly defined distinction in card-sorting speeds between good and poor conservers would have strengthened support for our hypotheses. Fig. 2 makes it clear, however, that the group which we operationally defined as poor conservers – i.e. subjects who, though clearly competent conservers in the traditional Piagetian test situations, failed at an early stage in solving more complex conservation problems – was consistently slower than the group of good conservers, albeit non-significantly except for 1·985 'bits'.

In general, there seems to be more support for the two stronger of our four hypotheses: that the difference in processing capacity should become more evident the greater the processing demands of the experimental task. This prediction was more substantially supported in the coding than in the card-sorting task where good conservers demonstrated a fairly regularly increasing superiority over the other groups. Exceptions to this were present on each occasion when four primary stimuli required recoding. We are of the opinion that these tasks may have been easier than tasks with an odd number of primary stimuli, because they always contained more clearly observable differences in the balance of colour symbols which facilitated the correct application of the rules for response selection (i.e. for 4, 12 and 20 coding steps). Nevertheless, the progressive dominance of the good conservers remained unaffected. In this experiment one might have wished for more consistent evidence of superiority of poor conservers over non-conservers, but possibly only on the dubious assumption of a large shared variance between handling redundancy and handling recoding operations. Apart from the difference in the two dependent variables, the cognitive operations cannot be considered, and were not assumed to be identical. Thus only 15 per cent of the correlations between the card-sorting and coding variables was significant.

The support for our predictions strengthens a suggestion in a previous article in this journal (Hamilton, 1972*b*), recently made with somewhat greater conviction (Hamilton & Moss, 1974), that the conserving operation may not be able to appear optimally if the capacity to process the information contained in a test display exceeds the capacity of the cognitive system. This implies a controversial assumption: that conservation of quantity is not an either/or, but rather a continuous variable. Evidence from three separate studies in the present series (Hamilton, 1972*b*; Hamilton & Moss, 1974; Hamilton, 1975), now supports this assumption. Success in the traditional Piagetian, concrete conservation tasks (involving a minimum of manipulation, and thus a small number of equational operations or steps), does not in any way predict success in conservation tasks containing a complex and extended manipulation of quantities, and thus a large number of processing steps. It is evident though that the present approach did not set out to explain the *emergence* of an invariance statement which has the power of logical necessity. This appears to be a separate, as yet, unsolved problem.

It is also evident that causal relationships have not been fully demonstrated. At the same time, more than correlational evidence is presented, if it is remembered that each level of redundancy (in Expt. I) and each level of recoding demands (in Expt. II), may be considered a separate experiment. In any case, a purely correlational approach would have demanded a larger sample of subjects with a better distribution of conservation scores, and not the criterion group strategy employed here. This strategy, however, has enabled us to show that criterial individual differences in the conserving skill, when one makes a deliberate attempt to demonstrate them, may be substantially dependent on the capacity to process sequences of logical steps which are both distinct and necessarily interdependent. Where the demands on a processing system exceed the present capacity of such a system, failure or errors in problem-solving are bound to occur. It is already known that maturational and experience factors determine the development of an individual's capacity to process increasing amounts of simultaneously available information. For this the increasing capacity of short-term, preliminary or temporary memory stores may be primarily responsible. Other factors, however, may also determine processing deficits even where they are developmentally unexpected (Hamilton, 1975), because of the intrusion of task-irrelevant information. The intrusion may be the result of what has been loosely termed 'test anxiety', which may be assumed to originate in socialization experiences interpreted as aversive, painful or threatening by children. Whatever the origin and quality of the task-irrelevant cognitive intrusion, it would appear to require definition as information competing for space and time in a limited capacity system.

Our definition of what constitutes information in conservation and coding tasks may seem excessively concrete and idiosyncratic. It appears to have some merit, however, by virtue of its objectivity in an area of work where clarity of concepts and processes is still somewhat deficient.

The experiments were supported by SSRC Grant HR 1556. We wish to thank the Council, our colleague Margaret Moss, and the Headmasters and Staff of local schools for their assistance.

REFERENCES

Crossman, E. R. F. W. (1953). Entropy and choice time: the effect of frequency unbalance on choice response. *Q. Jl exp. Psychol.* **5**, 41–51.

Hamilton, V. (1963). Size constancy and cue responsiveness in psychosis. *Br. J. Psychol.* **54**, 25–39.

Hamilton, V. (1966). Deficits in primitive perceptual and thinking skills in schizophrenia. *Nature, Lond.* **211**, 389–392.

Hamilton, V. (1971). Effect of maternal attitude on development of logical operations. *Percept. mot. Skills* **33**, 63–69.

Hamilton, V. (1972*a*). The size-constancy problem in schizophrenia: a cognitive skill analysis. *Br. J. Psychol.* **63**, 73–84.

Hamilton, V. (1972*b*). Continuities and individual differences in conservation. *Br. J. Psychol.* **63**, 429–440.

Hamilton, V. (1972*c*). Maternal rejection and conservation: an analysis of sub-optimal cognition. *J. Child Psychol. Psychiat.* **13**, 147–166.

Hamilton, V. (1975). Socialization anxiety and information processing: A capacity model of anxiety-induced performance deficits. In Sarason, I. G. & Spielberger, C. D. (eds), *Anxiety and Stress*, vol. 2. Washington, D.C.: Hemisphere Publ. Co.

Hamilton, V. & Moss, M. (1974). A method of scaling conservation of quantity problems by information content. *Child Dev.* **45**, 737–745.

Miller, G. A. (1956). The magical number seven, plus or minus two. *Psychol. Rev.* **63**, 81–97.

Newell, A. & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Osler, S. F. & Kofsky, E. (1965). Stimulus uncertainty as a variable in the development of conceptual ability. *J. exp. Child Psychol.* **2**, 264–279.

(*Manuscript received* 21 *November* 1974; *revised manuscript received* 15 *May* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 203–211
Printed in Great Britain

REPRODUCTION OF KOHS-TYPE FIGURES BY GHANAIAN CHILDREN: ORIENTATION ERRORS REVISITED

BY GUSTAV JAHODA
University of Strathclyde

First, a relationship was postulated between the experienced difficulty of patterns and the amount of disorientation of their reproductions. Secondly, while the major determinants of gross orientation errors among Africans seem to have been established, this in itself fails to explain why Europeans are little affected by them. It was proposed that the underlying factor might be the development of the concept of horizontality, which reduces dependence on the perceptual pull of the prevailing spatial framework.

These notions were tested with 44 Ghanaian schoolboys aged between 10 and 16. The first relationship was clearly confirmed. As for the second, there was some association between horizontality and rotation errors, but it failed to account for a major part of the variance. *Ex post facto* exploration of the data strongly suggests that degree of psychological differentiation is likely to be the key factor influencing rotation error.

In an early study primarily concerned with the assessment of abstract behaviour (Jahoda, 1956), it was reported that subjects paid little attention to the spatial orientation of blocks on the table, so that there was nearly always some rotation. It was suggested that subjects did not regard correct figure–ground relationships as part of their task and therefore paid little attention to them. Such a view implies that rotation errors are purely random in character, and this was disproved by the work of Serpell (1971*a*, *b*) and Deregowski (1972). They demonstrated that rotation errors are influenced by tendencies towards 'stability', i.e. the alignment of sides of a square parallel to the table edge. Serpell (1971*a*) in addition invoked the role of a 'focal point' previously discussed by Ghent (1961) and Strang (1967). This refers to an irregularity in a figure's contour (such as the angle of a V) leading to a tendency to locate this part at the top. Serpell himself regards this as a less important determinant, and there is American work indicating that focal point dependency is weaker among black as compared with white children (Barabasz, Dodd, Smith & Carter, 1970). Deregowski (1972) also showed that rotations occur in the direction of symmetry about the median plane, which may be related to focal point dependency; but there is no evidence concerning this relationship. The aim of the present study is to pursue two aspects of the general problem: one is to reopen the issue of attention on a minor key, the other to probe some possible factors underlying the disposition to rotate.

The starting point for the first was merely the personal impression gained by the writer in the course of fieldwork that lack of attention does play some part in the responses, especially where the patterns to be reproduced are difficult. Serpell's findings on this issue appeared to be negative, and he stated that 'systematic increases in the number of variable dimensions in a simple copying task did not affect the incidence of rotation errors by Zambian children' (1969, 1971*a*, p. 315). However, since his task was sequential rather than simultaneous, his evidence does not exclude the possibility of internal complexity having some effect. Moreover, Serpell distinguishes between 'rotation errors' that are rather gross and minor 'tremors' or 'disorientations', and the latter may well be a function of lack of attention without affecting his main contention. On this basis it was expected that where the internal structure of the patterns to be copied was relatively complex and thus more difficult, subjects would be less concerned about orientation. As the difficulty level may subjectively vary across individuals, it was operationally defined as the time taken for perceived completion. It was thus predicted that difficulty level, as defined, would be related to the amount of disorientation.

The second issue concerns the nature of the determinants of gross orientation errors. Deregowski calls them 'lawful' and Serpell 'preferences'; but why do these laws or preferences apply to African children far more than, say, Icelandic ones which Deregowski used as his comparison sample? It is likely that they are related to the difficulties experienced by Africans with certain aspects of space perception that have often been noted (cf. Jahoda, 1971). The tendency towards 'stability', when analysed, implies an inability to detach oneself from a given perceptual frame of reference. Piaget & Inhelder (1956) describe precisely such a problem in their discussion of projective space: 'There is a third case, however, where the difficulty proves insurmountable. This is when the line which the child has to imagine and construct conflicts with perceived curved or straight lines lying adjacent, such as on the "ground" offered by the table top (in the perceptual sense of "ground" as opposed to "figure"). In this case, imagining the line no longer consists of merely imitating a past or present perception, but entails creating new relationships within an existing pattern distinct from those sought after' (p. 163). The suggestion which emerges from this is that the subjects make gross orientation errors because they have not developed an adequate geometrical reference system.

There is a study carried out by Dadsetan in Geneva (described in Piaget, 1969, pp. 173 ff.) which bears directly on this issue. He showed that children's accuracy in judging the orientation of lines drawn inside tilted squares was related to their ability to draw the water level in inclined jars – in other words their concept of horizontality. This concept, studied by Dasen (1974) among Australian Aborigines, requires freedom from the constraining influences of perceptual configurations. This leads to the expectation that orientation errors of a gross kind would be inversely related to the degree of development of the concept of horizontality.

There is one other point that ought to be mentioned at the outset, since it will loom large in the subsequent discussion. The pilot work revealed something previously noted with Scottish children but altogether absent during the original Ghanaian (1956) study: namely a tendency towards 'final adjustment'. By this is meant that after completion of a particular reproduction subjects would deliberately grasp the whole set of blocks or shapes and rotate them into alignment with the table edge. It was therefore decided to take careful note of such behaviour in the course of the experiment.

METHOD

Subjects

These were obtained from the Staff Village attached to the University of Ghana at Legon. This village houses junior workers originating from all over the country, their tribal affiliations being mainly Ga, Akan, Ewe and some from Northern Ghana. They cover a variety of occupa-

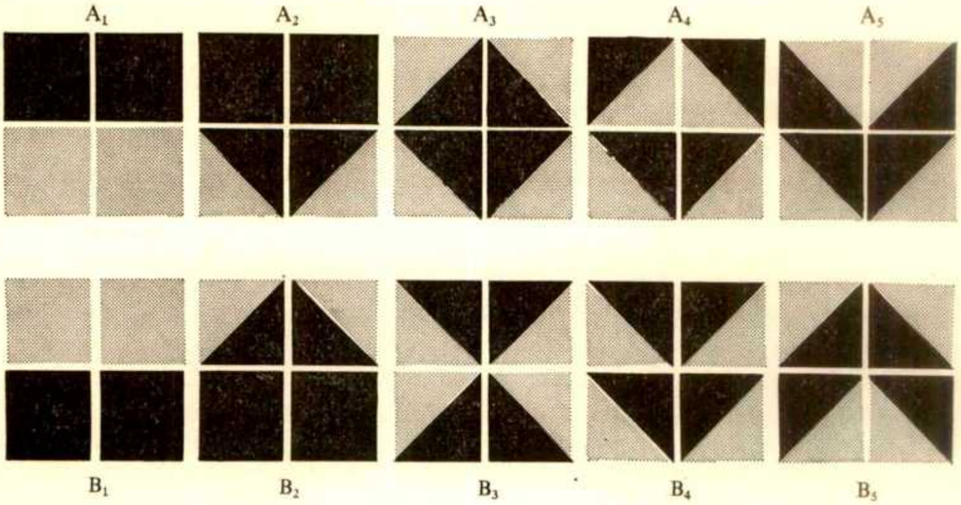


Fig. 1. Stimulus pattern sets A and B (plain black areas represent red, dotted ones blue).

tions including clerks, porters, stewards and labourers. Since the experiment was conducted during the Christmas holidays, subjects were recruited from among boys attending the primary and middle schools in the village through personal contacts. The 44 boys ranged in age from approximately 10 to 16 years, with a mean of about 13 years. Their classes included Primary 2–6 and Middle 1–4. The relationship between age and class attended is not very close, since starting ages are apt to vary and some pupils repeat classes. Subjects were rewarded for their help with soft drinks and biscuits.

Materials

The patterns used are shown in Fig. 1, A being the first series and B the second one. It will be noted that both sets are stable and symmetrical, thereby eliminating these sources of rotation error. The two sets were kept as similar as possible for the purpose of comparison in testing the first prediction concerning difficulty level. Thus it will be seen that B_1, B_2 and B_5 are simply upside-down versions of the corresponding A items, the remainder being of a similar order of complexity. Keeping the two sets as squares made it necessary to manipulate the spatial framework externally when assessing the effect of the concept of horizontality; and this was done by providing a diamond-shaped background. Thus, as will be further explained later, the white cardboard appeared as a square parallel to the table edge during the first set of trials; and it became a diamond with its sides at 45° to the table edge for the second set.

The size of each model card showing the patterns to be reproduced was 120 mm square, and the small shapes forming the patterns were 30 mm square. The response material consisted of plastic tiles of the same dimension as the models, i.e. 30 mm square. They were either red, blue, or divided diagonally into red and blue halves. For all trials four tiles were needed, and the appropriate ones were placed on the table before the start of each trial. A guide frame 100 mm square made of transparent plastic was also provided; it had an opening just over 60 mm square in the centre, so that the four tiles just fitted into it smoothly.

The material for the test of horizontality consisted of a spherical flask 150 mm high, half filled with blue-coloured water; an opaque cloth bag could be fitted over the flask. There were also duplicated responses sheets with outline drawings of empty flasks in six different positions (for further details see Dasen, 1970).

Procedure

On one side of a table a white piece of cardboard 200 mm square was fixed, its bottom side exactly aligned with the table edge. The subject was seated facing this, and was told he would have to copy some pictures after being shown how to do it. A training card with a simple pattern

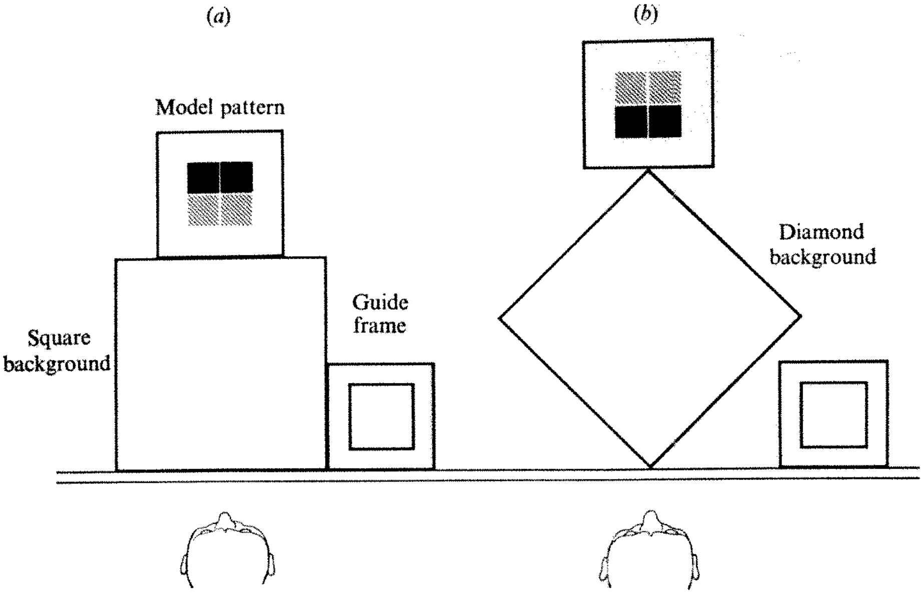


Fig. 2. Positioning of model card, background and guide frame; (*a*) layout for set A; (*b*) layout for set B.

(same as fig. 1, 4*d* in Deregowski (1972), p. 285) was placed exactly along the top edge of the white square and the guide frame put in its centre; the frame was then adjusted with deliberate slowness and care until it was straight, i.e. parallel to the model card and the table edge. The experimenter then constructed the pattern inside the guide frame. Following this demonstration, the subject was asked to perform the task, stressing that his picture should be in every way the same as the model; he was also instructed to compare his construction with the model before saying that he had finished. All subjects managed to do the training task correctly, but many left the guide frame in varying degrees off the horizontal; in those cases the experimenter adjusted it, telling them to remember that their picture must be the same in every way.

Thereafter the first set of trials began. Before each, the experimenter positioned the guide frame along the side of the white square and exactly parallel to it; the general layout is shown in Fig. 2 (*a*). The four tiles required were scattered randomly on the table to the right of the guide frame. It was ensured that the subject always started by positioning the guide frame and constructed the pattern within it. The time elapsed from the moment the subject took up the guide frame to reported completion was recorded to the nearest second. The assessment of orientation errors presented a problem; it had been hoped to photograph the responses as Deregowski (1972) had done, but in the event the necessary facilities proved unavailable. Therefore it was decided, after trying it out with a pilot group, simply to measure each response with a large circular protractor. Pilot subjects questioned about this gave hardly any indication of familiarity with the instrument or its purpose; and while there might possibly have been some effect in reducing the absolute amount of rotation, though this is considered unlikely, it could not have affected the relationships to be tested. The main drawback was that it was very time consuming and entailed the postponement of horizontality concept assessment to the final stage in order to guard against possible experimenter bias. Thus after the first series of trials was over subjects were otherwise occupied for a brief interval and then sat down again. The procedure for the second series was exactly the same as for the first, except that the white cardboard background was now oriented in the shape of a diamond, whose lowest point just touched the table edge; both the model card and guide frame continued to be aligned with the table edge as illustrated in Fig. 2 (*b*). Subjects were merely told that their task was just the same as before, making a picture that is in every way the same as the model.

When the subjects reported completion of a particular item, the responses of some were occasionally incorrect. The most common types of errors maintained symmetry about the vertical axis, and there was no problem about measuring rotation. Yet there were other types of error which rendered the pattern completely ambiguous in terms of orientation. In those cases the deviation of the base line from the horizontal was assessed; this will have meant frequent underestimation of the amount of rotation, but in terms of the relationships postulated it is a conservative procedure.

The behaviour of subjects at the end of each trial was closely scrutinized for 'final adjustment', i.e. rotating the guide frame into proper alignment with the horizontal. The frequency of adjustments per subject ranged from a mode of none to one subject who adjusted after each of the ten trials, the mean being 2·3. A large majority of these adjustments were highly effective: on more than half such trials the deviation from the horizontal was too small to be accurately measured, and in over four-fifths of the trials it remained within 5°. Subsequently subjects were divided at the median and categorized as 'low' (zero or one) versus 'high' (more than two) adjusters.

The final part of the procedure was the test for the concept of horizontality, which followed precisely the first stage as described in detail by Dasen (1970). It should be pointed out that this was a considerably shortened version of the test, but it served to give a rough indication of the extent to which the concept had been acquired. Results were scored by assigning one point for each correct drawing, so that scores ranged from 1 to 6. The mean score was 3·07, but the distribution tended towards bimodality with peaks at 1 and 4; only two subjects achieved the maximum score. For the purpose of analysis subjects were dichotomized into 'low horizontals' whose score fell between 1 and 3, and high ones scoring more than 3.

RESULTS

Disorientation in relation to difficulty

The hypothesis is that the greater the difficulty of the task, as indicated by completion time, the greater will be the deviations from the horizontal. The mode of analysis was to take for each subject the mean of the two shortest and two longest trials, i.e. omitting the intermediate one; where there was a tie, random allocation was used. Since any adjustments would distort the relationship, subjects who made more than one adjustment were excluded. This analysis was straightforward for set A, but not for set B. It will be recalled that an additional variable was involved here, namely the diamond-shaped background which tends to elicit rotation errors (as distinct from disorientations) towards 45 degrees. In order to cope with this, measures of displacement were taken from whichever reference axis (0° or 45°) was closer. This yields a purified measure of disorientation, uncontaminated by large deviation scores over 22°. The results for both sets A and B are given in Table 1.

It will be seen from this that there is a highly significant association (Wilcoxon matched-pairs signed-ranks) of the kind predicted. Moreover, in spite of the 'purification', the overall deviation scores are higher for set B. In order to reduce the likelihood that this might merely reflect greater inherent difficulty of this set, the difference was tested using only the results for items 1, 2 and 5 which are identical but for the fact that one set is a 180° inversion of the other. On this basis the difference is significant (P less than 0·01), indicating that the diamond-shaped background had a generalized and independent effect in increasing the amount of disorientation.

Factors influencing rotation errors

The prior expectation was that subjects with a more highly developed concept of horizontality would be *less* likely to make rotation errors when faced with a conflicting frame of reference, i.e. the diamond-shaped background. It should be noted that the manner of determining rotation errors was quite different from that of assessing

Table 1. *Mean deviation scores for pairs of slowest and fastest trials*

		Longest mean		Shortest mean		
	n	Time (sec)	Rotation (deg)	Time (sec)	Rotation (deg)	Rotation difference P less than
Series A_1–A_5	27	38·65	7·97	16 61	4·06	0·001
Series B_1–B_5	31	45·19	14·37	13·24	6·58	0·001

Table 2. *Rotation errors in relation to horizontality and adjustments*

		Horizontality		
		Low	High	Total
Rotation errors	Present	13	5	18
	Absent	11	15	26
		24	20	44

$P < 0{\cdot}05$ (one-tailed)

		Adjustments		
		Low	High	Total
Rotation errors	Present	14	4	18
	Absent	10	16	26
		24	20	44

$P < 0{\cdot}025$ (two-tailed)

'disorientation'. The base line for rotation errors was the horizontal, and a rotation error defined on any one trial as a deviation from the horizontal exceeding 22°. Subjects were subsequently classified as being in the 'rotation error' category if their deviation scores were larger than 22° on two or more of the five set B trials.* For this part of the analysis the data were cast in the form of 2 × 2 contingency tables to which Fisher's exact test was applied.

The outcome is shown at the top of Table 2. Although the relationship found was in the expected direction, it is rather weak. Hence it was decided to conduct some further exploration, beginning by looking at adjustments. From the bottom part of Table 2 it will be seen that this variable appears to be more strongly associated with rotation errors, which suggested the need to find out more about its nature. Since this had to be done *post hoc*, the only way open was to examine the general pattern of relationships among all the variables on which information was available.

The tool employed for this purpose was Kendall's (1970) tau_b, a coefficient measuring association in a 2 × 2 table. It is important to note that the entries into such a table consist of the number of subjects who possess or do not possess certain attributes; in other words it does *not* relate to performance on particular trials as would parametric correlation, which is not applicable in the present case. On each of the

---

* A problem was presented by an unusual subject high on horizontality who had negligible deviation scores on set A; he was the single one who made adjustments after each trial and obtained the maximum score of 10. This subject, remarking 'I shall do it this way', proceeded during set B to align his responses very accurately to the 45° background. In his particular case, which was the only one of its kind, this had clearly been a quite deliberate choice. Nevertheless it was decided to classify him as a 'rotation error' subject so as to avoid possible bias favouring the hypothesis.

Table 3. *Tau$_b$ coefficient matrix for main variables*

	2	3	4	5	6
1. Age	0·76 (0·001)	0·22	0·08	−0·03	−0·06
2. School class	—	0·34 (0·03)	0·19	−0·21	−0·14
3. Horizontality	—	—	0·36 (0·02)	−0·30 (0·05)	−0·26
4. Adjustment	—	—	—	−0·39 (0·01)	−0·45 (0·004)
5. Rotation error	—	—	—	—	0·32 (0·03)
6. Pattern reproduction error					

relevant variables subjects were dichotomized at or near the median, as shown below:

Age: below 14 vs. 14 and above

School class: primary vs. middle school

Horizontality: score below 4 vs. 4 or more

Adjustment: less than 2 vs. 2 or more

Rotation error: less than 2 vs. 2 or more

Pattern reproduction error: none vs. 1 or more

On this basis tau$_b$ coefficients were then computed, and as an indication of the existence (not the magnitude) of a given association the one-tailed probability value of the corresponding contingency table was determined. The outcome is presented in Table 3, where probability values are added in parentheses; absence of brackets means a non-significant association. While signs are essentially arbitrary in a matrix of this kind, they have been added in an effort to assist the understanding of the relationships. Thus it will be clear that there is a regular inverse relationship between the first four and the last two variables, e.g. subjects high on horizontality tend to make fewer rotation and reproduction errors.

In spite of the crudity of this procedure, some interesting patterns may be discerned within the matrix. In particular, three clusters will be singled out, going from top left to bottom right:

(*a*) age, school class and horizontality;
(*b*) horizontality, adjustment and rotation error;
(*c*) adjustment, rotation error and pattern error.

Now (*a*) would suggest that for this sample the development of the concept of horizontality may be largely a function of formal schooling. Cluster (*b*) merely adds to the information in Table 2 that horizontality and adjustment are significantly related. But what appears to be the key to the whole puzzle is contained in (*c*), where the largest single coefficient occurs (apart from the trivial age × school class); namely the inverse relationship between adjustment and errors in pattern completion. Now block design is of course precisely the task employed in many cross-cultural studies in order to assess psychological differentiation (Witkin & Berry, 1975). It would seem reasonable to infer, therefore, that the common factor underlying both low pattern errors and high adjustments is likely to be field-independence. Such an interpretation also fits in with other aspects of the matrix, in as much as both variables exhibit the kind of modest association with age and education that would be expected; in other words, neither appears to be predominantly a function of age and/or amount of education.*

DISCUSSION

The first part of this study clearly confirms the expectation that subjective difficulty influences the extent of minor deviations from horizontal alignments in pattern reproduction tasks. There appear to be no cross-cultural comparative studies of such disorientations, but it is the writer's impression that they are far more pronounced among African children. This would be consonant with the fact that African children generally have more difficulty with this type of spatial–perceptual task. However, the relationship should also apply to European children, though no evidence is at present available.

The second part of the study, building on the work of Serpell and Deregowski, started off with the idea that the degree of development of the concept of horizontality might account for variations in the extent to which African children rotated designs, as well as explaining differences between them and European children. The findings lent some support to this view, but it became clear that grasp of the concept of horizontality accounted for only part of the variance in the behaviour displayed. Further unplanned exploration of the available data led to the conclusion that degree of psychological differentiation is likely to be the major underlying determinant. On reflexion this seems to make good sense; for while an individual may have a reasonably developed concept of horizontality, it does not necessarily follow that he is bound to make appropriate use of it in particular situations – he may still yield to the purely perceptual pull of an immediate and dominant spatial frame of reference. This is especially liable to occur when the spatial configuration is apprehended as an undifferentiated whole, as is characteristic of field-dependent persons. On the other hand a subject who has a clearly articulated view of the relevant components of the visual–spatial field, i.e. model – immediate background (diamond shape) – broader background (table), is better equipped to abstract from the immediate background and maintain the orientation of the reproduction the same as that of the model.

This notion of the role of psychological differentiation emerged from a scrutiny of the results, long after the material was collected. Hence it could only be tested with the data at hand, which were rather crude instruments for the purpose. Subsequent search of the literature yielded only one relevant study by Blum & Chagnon

---

* One reviewer suggested an alternative interpretation, namely that the association between pattern error, adjustment error and rotation error could be mediated by the structure of the task. The argument is that subjects whose reproductions contain pattern errors would be less likely to perceive that the whole design is rotated and needs adjustment, since the recognition of the orientation differences requires that some parts of the two figures can be matched. This presupposes that most pattern errors destroyed the symmetry about the vertical axis which, as already mentioned above, was not the case. Moreover, it seems to imply a strong association between pattern and rotation errors *at the level of particular trials*; and this is quite distinct from the association appearing in Table 3. When the data are analysed at this level, taking all trials in which one or other error occurs and expressing this as a percentage of total number of relevant trials, the outcome is as follows: rotation errors *without* pattern errors – 48·3 per cent, pattern errors *without* rotation errors – 32·6 per cent, both pattern *and* rotation errors – 19·1 per cent. It will be evident that the conjunction of the two error types is relatively infrequent, and the alternative interpretation is therefore unlikely to hold.

(1967), which found an association between rotation scores and cultural norms theoretically expected to affect psychological differentiation. Since extent of field-dependence was inferred rather than specifically measured, their results provide only rather weak and indirect support for the interpretation offered here. Thus while the internal evidence obtained appears coherent, and consistent with most of the available information concerning both orientation errors and psychological differentiation, adequate confirmation requires research directly aimed at testing the hypothesis emerging from the present study.

Thanks are due to Mr A. Afrifah for his assistance in conducting the study. The writer is greatly indebted to Drs Deregowski and Serpell for their helpful comments. The work is part of a cross-cultural project supported by the Social Science Research Council.

References

Barabasz, A. F., Dodd, J. M., Smith, M. & Carter, D. E. (1970). Focal-point dependency in inversion perception among Negro, urban Caucasian and rural Caucasian children. *Percept. mot. Skills* **31**, 136–138.

Blum, D. M. & Chagnon, J. G. (1967). The effect of age, sex and language on rotation in a visual-motor task. *J. soc. Psychol.* **71**, 125–132.

Dasen, P. (1970). Cognitive development in aborigines of Central Australia. Unpublished PhD thesis, Australian National University, Canberra.

Dasen, P. (1974). The influence of ecology, culture and European contact on cognitive development in Australian Aborigines. In J. W. Berry & P. Dasen (eds), *Culture and Cognition: Readings in Cross-Cultural Psychology*. London: Methuen.

Deregowski, J. B. (1972). Reproduction of orientation of Kohs-type figures: a cross-cultural study. *Br. J. Psychol.* **63**, 283–296.

Ghent, L. (1961). Recognition by children of realistic figures presented in various orientations. *Am. J. Psychol.* **74**, 177–190.

Jahoda, G. (1956). Assessment of abstract behaviour on a non-western society. *J. abnorm. soc. Psychol.* **53**, 237–243.

Jahoda, G. (1971). Retinal pigmentation, illusion susceptibility and space perception. *Int. J. Psychol.* **6**, 199–208.

Kendall, M. G. (1970). *Rank Correlation Methods*. London: Griffin.

Piaget, J. (1969). *The Mechanisms of Perception*. London: Routledge & Kegan Paul.

Piaget, J. & Inhelder, B. (1956). *The Child's Conception of Space*. London: Routledge & Kegan Paul.

Serpell, R. (1969). Cross-cultural differences in the difficulty of copying orientation. *H.D.R.U. Reports* **12**.

Serpell, R. (1971*a*). Discrimination of orientation by Zambian children. *J. comp. physiol. Psychol.* **75**, 312–316.

Serpell, R. (1971*b*). Preference for specific orientation of abstract shapes among Zambian children. *J. Cross-Cultural Psychol.* **2**, 225–239.

Strang, H. (1967). Relationship between focal point location and inversion perception at three age levels. *J. genet. Psychol.* **111**, 3–8.

Witkin, H. & Berry, J. W. (1975). Psychological differentiation in cross-cultural perspective *J. Cross-Cultural Psychol.* **6**, 4–87.

(*Manuscript received* 15 *March* 1975; *revised manuscript received* 27 *May* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 213–230
Printed in Great Britain

WERE CATTELL'S 'PERSONALITY SPHERE' FACTORS CORRECTLY IDENTIFIED IN THE FIRST INSTANCE?

BY EDGAR HOWARTH
University of Alberta, Canada

Cattell's system of personality factors has not changed in essentials since his foundation 'personality sphere' studies of 30 years ago, although the original 12 factor (A through L) system has been extended and has found worldwide application in the Sixteen Personality Factor Questionnaire (16 PF). The article begins with a summary of recent criticisms of the latter, as reflecting upon Cattell's system, and then proceeds to a specific re-examination of the foundations of that system in order to put the question: Were Cattell's 'personality sphere' factors correctly identified in the first instance? Instead of 12 factors, it is herein suggested that possibly six clear-cut factors might have been extracted originally, factors which accord well with more recent work carried out independently from Cattell's laboratory. The implications of this are examined, accompanied with recommendations for more accurate delineation of personality factors in the future.

Cattell's long continued and widely documented research into personality factors represents the single most monumental effort to operationalize those sources of individual differences often referred to as 'temperament' or 'personality' (and is comparable to that of Guilford on cognitive factors). According to Berg (1972) 'Cattell's major contribution will undoubtedly be seen as the discovery of the major factorial dimensions in rating, questionnaire and objective tests'. It is this system with which we are now concerned, and in order to clear up any initial misunderstanding it should be stated that this writer does not belong to that large camp of North American psychologists who are either indifferent or hostile to Cattell's work [in marked contrast to Europe – see Berg (1972) for a fuller discussion of the matter]. Rather, the present article *is* concerned with the accurate delineation of personality factors and, quite clearly, had Cattell already outlined or charted the 'personality sphere', then subsequent investigators could safely use this chart while exploring the far corners of the sphere. Moreover, in view both of the laudable scope and intention of Cattell's work (Howarth & Cattell, 1973) the present article is written with both regret and reservation, and only under the impetus of scientific truth and accuracy. Nor should a critical examination of a particular system be taken as reflecting upon multivariate methods (Adcock, 1972) *per se.*

The present article does not, therefore, call into question either the scope and intention of Cattell's work, which is unrivalled in both respects, but it *does* have the temerity to question certain methodological practices and assumptions on the part of Cattell. We find ourselves, therefore, somewhat in the role of David facing Goliath with a small army at our back and a large army in front! Nor do we question Cattell's strategy for these reasons: Fiske (1971) has recently reviewed the various modes of measuring personality, each with its devotees and limitations, and each not always easily relatable to other modes. However, it would seem that ratings might offer a more objective starting-point than questionnaires and Cattell's strategy was to begin with the former and to attempt to comprehensively combine a large number of these into a single factor analysis in order to explore the 'personality sphere'. His assumption was that the 18000 or so dictionary terms relating to temperament (and, more specifically, the 4000 or so general terms) could be more parsimoniously categorized. Having done so, he then proceeded to other measurement media, and suggested the principle of indifference of media (cf. Spearman's principle of indifference of indicator).

Despite the potential soundness of this strategy, it will be herein argued that, in Cattell's case, this was somewhat vitiated by poor tactics, in that (*a*) recent evidence has thrown doubt on the subsequently developed questionnaire factors – which were targeted to the original rating factors: (*b*) certain methodological assumptions formed an integral portion of those tactics. In short, we are going to argue the case that relatively few of his factors will long survive in the face of present-day technological advances, and we are going to (implicitly) regret the fact that he had not used certain superior facilities (e.g. ILLIAC) long available to him to update or correct his system.

Cattell's strategy was, thus, to begin with ratings (Cattell, 1946) and having identified 171 bipolar scales among 4000 or so general trait terms, he then obtained ratings of 100 subjects. Using Thurstone's approximation tables he then laid out a matrix of interrelations on a 14 ft square table at Harvard University, inspection of which revealed a smaller number of clusters among the variables and indicated the possibility for further simplification. A smaller set of representative scales was subsequently factored to provide the 12-factor system which he featured in his 1946 book.*

Having moved to the University of Illinois shortly thereafter he carried out a study with methodological improvements such as better rating procedures and using regular product-moment correlations, and thereby claimed (Cattell, 1947) to have replicated 9 out of 12 of the original factors and so felt that the time was ripe to embody these verified factors in questionnaire measurement.

Although the rating factors long remained the foundation of his system, Cattell (1973) has moved to the position that the questionnaire mode is the main exemplar of his system, and claims that there are at least 12 and probably 16–21 rating factors matching the main Q series. None the less, as recently as 1970 he specifically stated, as he had done in previous manuals, that the Q factors were targeted to the rating factors (Cattell, Eber & Tatsuoka, 1970).

We will now survey some of the recent evidence which has accumulated to indicate that Cattell's continued building on the foundation studies contains some over-hasty methodological assumptions. For example, his work on ratings has been scrutinized by Tupes & Christal (1961) and Norman (1969) and these more recent careful investigations have reduced Cattell's claimed 12-factor system down to a more modest 'recurrent five-factor structure'. Digman's (1972) research has also pointed to the conclusion that there were fewer factors in ratings than previously suggested by

---

* During a visit to Europe shortly after World War II, Cattell met Kretschmer and told him of his factor work, in particular that he had found a leading factor which he had named 'cyclothymia–schizothymia'. On hearing this, Kretschmer remarked, 'Ach, Herr Cattell, *now* I believe in factor analysis'.

Cattell; indeed, the eigenvalues* from this work indicated seven factors, as from a large sample of 456 children, rated on 49 scales, these were:

14·3 8·6 3·0 1·9 1·5 1·2 1·1 0·94 0·80 0·79

While the 'number of factors' problems is by no means finally solved, it seems reasonable to suppose that Cattell's claims were, possibly, somewhat over-extended. Indeed, if only investigators could verify so many factors as he claimed, that would present a different basis for evaluation of his position. Let us, bearing these findings in mind, pursue the question: if Cattell's rating factors were over-extended, and in that sense, possibly inaccurate in some respects, how did this subsequently affect his targeting of the later-developed 16 PF (questionnaire) factors to the original rating factors?

Howarth & Browne (1971*a*) carried out a detailed item-factor (i.e. using each item as one variable, rather than groups or packages of items) analysis of the 16 PF. This entailed computing the coefficients among 187 variables (i.e. 17391 correlations, or if the first two and final buffer items are omitted, 16836 coefficients) followed by a factor analysis of the entire correlation matrix preceding analytic simple structure rotation and visual (Rotoplot – Cattell's own program) rotation to oblique simple structure. These authors found only 10 (not 16) clear factors in the 16 PF, factors which did not *in any way* support Cattell's views, but which, on the other hand, *did* agree with their previous attempt (Howarth & Browne, 1971*b*) to verify Sells' factors. Lest these two studies cited (the item-factor analysis of the 16 PF, and the quite independent study by these writers using marker items from Sells' study) are considered by some readers – by comparison to the widely known position of Cattell – to be isolated and somewhat idiosyncratic, further verification of the findings of Howarth & Brown *vis-à-vis* the 16 PF, has appeared in a manifestly independent study, based on the latest version of the 16 PF and New Zealand subject samples (Adcock, Adcock & Walkey, 1974). Moreover, Karson & O'Dell (1974) have also found that the majority of the primary factors, supposedly measured by this inventory, were unsatisfactory (they found that factors E, F, G, H were the least unsatisfactory). On the basis of these evidences it would begin to appear that a case could be put for re-evaluation.

Additional critical evidence has originated from independent investigators in the United States and United Kingdom, respectively. Sells, Demaree & Will (1970), having first obtained from each major investigator (Cattell and Guilford) items which best represented their factors, subsequently obtained responses from over 2000 subjects and factored a gigantic correlation matrix containing 179700 non-repeated coefficients. (There were inter-item correlations among the 600 marker items submitted, 300 being from Cattell, 300 from Guilford and representing the major portion of their factor systems.) This tremendous technical feat, as yet unsurpassed, analysed in one factor analysis the combined Cattell and Guilford markers in order to see which factors represented in their disparate systems might emerge from joint analysis. Although not especially favourable to either investigator, the results were least favourable to Cattell, so that the factors obtained have been vigorously criticized (see especially Cattell, 1973) mainly on the grounds that many of Sells' primary factors appeared to be second-order (more general) factor combinations. Howarth & Browne (1971*b*) selected the highest loading items from a group of the clearest factors in Sells' study and subjected these to a detailed factor analysis on a new subject sample with a view to testing whether these were primaries or secondary factors, and the results were taken to be quite favourable to Sells. In subsequent (and as yet unpublished) work the latter investigators, with the cooperation of 16 Canadian universities, carried out a factor analysis of 400 items, culled from 50 years of personality tests. The conclusion was that neither Sells nor Guilford nor Cattell were correct, but that their 'correctness' could be ordered in terms of the comprehensiveness of their item pools. In additional (and as yet unpublished) work, the writer has tested a variety of subject samples at different socioeconomic levels and in two language-cultures (so far) leading to his present view that the factors were definitely not as (for example) Cattell in particular proposed.

---

* Cattell's data for a sample of 133 male students, rated on 36 scales, provided the following eigenvalues (unrotated solution from the article): 7·7, 5·3, 3·2, 1·4, 1·0, 0·85, 0·71, 0·71. Readers familiar with the Kaiser–Guttman criterion will be able to draw their own conclusions from the above (the eigenvalues were not presented in the original report) and compare this with Cattell's most recent comment on the matter (Cattell, 1973) – 'even more quixotic than this button-pushing is the whimsical choice of a number of factors different from that dictated by any statistical test' (p. 285).

A similar conclusion regarding Cattell's factors was arrived at by Eysenck in the United Kingdom (somewhat paralleling Sells' work in that he also used selected markers submitted by Cattell and Guilford) (Eysenck & Eysenck, 1969). Because of local computer limitations these were not, and never have been, factored in a single analysis. However, their three separate factor analyses – of Eysenck, Cattell and Guilford item pools, respectively – are of great interest, and the Eysencks stated forthrightly that 'the primary factors which have been unearthed and used in personality research, such as the Cattell and Guilford ones, fail because they are not in fact unitary, univocal combinations of items measuring the same fundamental trait; instead, they are half arbitrary combinations of items sharing functional equivalence only to a limited degree and, thus, quite unfitted to bear the burden imposed upon them by their originators' (*op. cit.* p. 231).

In addition to these technical analyses of item pools, and the before-mentioned analyses of the factors in the 16 PF, there are other kinds of analyses in the literature which may be even more convincing to the non-factor-analytically inclined (but otherwise psychometrically competent) reader. Howarth, Browne & Marceau (1972) item-analysed the 16 PF and calculated (among other things) proportions of significant relationships of items both with their own assigned factor scales and outside these scales. They reported an overall efficiency level of only 33·4 per cent for Cattell's factor scales, the best factors being Parmia H (76·9 per cent), Ergic tension Q4 (58·9 per cent), Surgency F (55·1 per cent), and Dominance E (46·2 per cent). Moreover, there was considerable agreement as to which of Cattell's scales were best and which were worst, with a previous item analysis carried out by Levonian (1961) using an earlier version of the questionnaire. Additionally, these results concerning the relative accuracies of Cattell's scales have been more recently confirmed by Adcock (1974) using the latest version of the 16 PF. Supporting these detailed item analyses Eysenck (1972) has analysed the interrelations between Cattell's scales and has commented that 'clearly his hypothesis of 16 functionally independent factors requires considerable support if it is to be continued being accepted by test users'. Also of relevance is the recent information on scale intercorrelations presented by Goldberg, Norman & Swartz (1972) with the comment that 'As has been pointed out repeatedly in reviews of the 16 PF a number of scales have higher between-item correlations with other scales than with those to which they are supposedly equivalents (p. 7). Additional criticism will be found in Greif's (1970) analysis of the German version of the questionnaire.

Thus five lines of evidence converge: (1) from rating studies; (2) from item-factoring; (3) from factoring item pools; (4) from item analyses; (5) from inter-scale correlations. These have been briefly presented for two reasons: (*a*) to contrast, for the reader, the wide renown of Cattell's factors with their recent critical examinations by independent and concerned investigators; (*b*) to indicate that we may now usefully examine the accuracy of the *foundations* on which the house of Cattell was erected.

THE REANALYSIS

We should bear in mind that 30 years ago, when Cattell carried out his pioneering studies, there were few guidelines for making certain technical decisions (e.g. as to the number of factors problem) and, therefore, he was perfectly at liberty to follow his own views; moreover, those methodological assumptions which therein appeared have since been repeated over and over in Cattell's work [for a recent defence and explanation on his part see pages 281–319 of Cattell (1973)].

(1) He abstracts a relatively large number of factors from a given data matrix, thus giving importance (possibly more than one would wish, in terms of replicability, scale accuracy, and so forth) to factors of lesser variance.

(2) Connected with the previous point, he has made use of quite small factor loadings (projections of a variable on a factor axis), and also of relatively small groups of variables said to load or 'mark' a factor. This is an absolutely vital matter for when one has completed the analysis one is faced with semantic and/or observational decisions as to factor meaning. If a factor is a hypothetical variable lying in the test space how many different putative aspects of it are needed?

(3) When criticized [e.g. see pages 281–319 of Cattell (1973)] Cattell has recourse to the fact that he is the world's most experienced, and skilled, visual rotator of factor solutions. Having been taught this art by Cattell himself, the writer would not disagree with this, but would only point out that it is, most definitely, an art, so that his view is that the objectivity of analytic methods somewhat justifies their current wide usage. Although by no means so experienced as Cattell, the writer has spent the last several years in comparing solutions obtained by such analytic programs as Varimax and Promax with visually rotated solutions and has been impressed by both the speed and general accuracy of the 'canned' programs. When one considers that in, say, 3 min one can obtain rotations which might take several weeks by visual rotation the reason for the wide displacement of the latter by the former seems clear.

This reanalysis of Cattell's foundation rating studies (via the more accurate of the two) was not, therefore, undertaken with Rotoplot, as was the writer's previous 16 PF analysis, for two reasons: (*a*) it would be possible to produce an infinite number of 'alternate' visuality rotated solutions to that of Cattell and he could, quite properly, question the comparative skill of the other rotator; (*b*) as the literature was recommending the Harris–Kaiser method and as the writer had been comparing the visual method with various analytic methods, he decided to follow a procedure which had developed over the course of several years among the members of his research group, and to combine this procedure with the 'flexible' (a term which will be illustrated below) Harris–Kaiser development of Hakstian (1970). Before discussing these technical developments further, it will be best to return to the methodology of Cattell in order to compare this with the present method.

Turning to Cattell (1947) we find certain methodological improvements in subject sampling, rating precautions and the calculation of coefficients, although the characteristic weaknesses (in this writer's opinion) remain, and we find him extracting a dozen or so factors from only 36 variables, using quite small salients for factor identification and so forth. This last point becomes quite crucial when we realize that his purpose in the study was to attempt to replicate (i.e. re-identify) the majority of the factors in the initial 12-factor system of the personality sphere. As this point is so important we will now pursue it further.

While an inspection of the tables in the original article would be desirable, we have no facility (i.e. space) for doing so here and will attempt to present the flavour of an argument in place of a detailed examination. Cattell used a rather relaxed criterion for determining which variables to 'pull-in' and to increase the size of an identificatory group of variables to establish factor meaning and to be able to claim replication in various instances. To be sure, as there is no agreed method, even at the present time, for determining the significance of a factor loading he was entitled to lower his values, but one has to bear in mind that he then, and has consistently since that time, adopted simple structure assumptions. In this connection (and the writer also follows this approach), as one uses loadings as small as 0·20 or even lower in Cattell's case, one begins inevitably to vitiate whatever simple structure there might be in the data. It is this writer's view that one should couple with the Thurstonian simple structure rotation methods (e.g. as used in Varimax, Promax, and across the spectrum of Hakstian H–K rotations) a much stricter acceptance level for a salient. As a working rule which has developed in the writer's research group, we have found – across various types and sizes of data matrix – that we prefer not to go below 0·35 for othogonal solutions. For oblique solutions, such as those featured in Cattell's work, a higher value would be safer, and, in fact, we generally do not go below 0·40 in the case of these solutions. The reason for this is that the 'pattern on primary' solutions inflate the coefficients. If one adopts stricter values as suggested here, then one is led to identify fewer factors, so that in the event Cattell's first two assumptions, or working methods, are interrelated (many factors, use of small loadings).

We will now illustrate this: When one examines his rotated factor solution in the 1947 article (also see that in the original Harvard study) one is struck by the fact that only *three* factors approach the following criterion – four or more salients having appreciable values, for an oblique solution 0·40. Moreover, if we relax this somewhat and accept 0·35 (which one is loath to do in terms of one's experience) then, still, only *five* factors qualify, the weakest of these possessing only three salients each; insufficient to decide on factor meaning. These best five factors (and the reader may wish to compare this finding with the best factors found in various questionnaire studies as outlined above) were: Dominance E, Superego G, Parmia H, Surgency F, plus a fifth factor tentatively identified by Cattell 'as having a slight resemblance to Factor D of the original research, but the overall resemblance is undoubtedly to Factor I' (1947, p. 215). This fifth factor, we note, had only three salients greater than 0·35. Overall, and by the methods and assumptions we have stated, Cattell claimed to have recovered 9 out of 12 of the Harvard factors, and thereby to have confirmed the major portion, on a different subject sample and with certain methodological differences in the rating procedures and so forth, of the main factors in the personality sphere. Most certainly, to have defined personality so comprehensively and definitively would have been a major breakthrough, and this was Cattell's intention, to achieve which he had first surveyed 17 previous (partial) studies and had then attempted more comprehensive factorial investigation, aiming at a fuller taxonomy. This need, as we will illustrate both in terms of what factors Cattell might have reported, and in terms of the final discussion of what is needed to achieve this goal, still pertains at the present time – 30 years after Cattell's pioneering efforts.

METHOD

The rotation procedure applied to the unrotated factor solution in Cattell (1947) was originated by Harris & Kaiser (1964), initially programmed at the University of Wisconsin and subsequently developed into the form used herein by Hakstian (1970). In the last 3 years the writer has become familiar with the advantages of this program, and has coupled it with other procedures now to be outlined. The program is more flexible than the usual 'canned' Varimax or Promax combination (the latter being based on the former) in that it permits standard solutions such as Varimax, plus standard H–K solutions (independent cluster; $P'P$ proportional to phi) but adds to these the following facility which is most useful for exploratory work using the methods of the present report:

It is conceivable that values exist for P between 0 (independent cluster) and 1 (orthogonal solution) that result in superior oblique solutions for certain kinds of data. In particular, values of P between 0·5 and 1 may yield superior solutions (and more nearly orthogonal) for particularly complex, many factor data. The present program is a modification of the University of Wisconsin Harris–Kaiser, the earlier program allowing for only two values of P (0; 0·5) and two methods of orthogonal rotation (Quartimax, Varimax). The present program yields a large number of Harris–Kaiser solutions by permitting all feasible values for P and W. Additionally, all possible orthomax-orthogonal solutions are obtainable by setting P equal to unity (Hakstian, 1971, p. 703).

Although the potentialities of this newly developed program can hardly be said to have been fully explored, the earlier, more restricted version has already been shown to be at least the equal of currently popular methods such as Varimax and Promax in comparative trials by Dielman, Cattell & Wagner (1972). It was, indeed, because of the recommendations of the latter authors that we felt it to be both appropriate and timely for us to apply the Hakstian H–K, in combination with those procedural steps and precautions which we had found to be viable, to Cattell's data. Because of the accumulating evidence (see above), we felt that there was a case for believing that a more restricted (but more robust) set of factors might be found in Cattell's data, but the spirit of the inquiry was quite empirical in that any decisions were left to the analytic method without the interpositions of any prior assumptions as to 'number of factors', although there was one point on which this investigator refused to hedge, which concerned the minimal values for accepting factor salients *qua* salients. We also bore in mind Hakstian's, recommendation that superior solutions were likely to emerge not only with the intermediate P values but with the W parameter set to half the number of expected factors (e.g. with ten factors 'expected' we set the latter parameter to five in certain cases).

RESULTS

Ten-factor solutions

The 36 chosen scales in Cattell (1947) represented condensation over 4000 personality descriptors by the methods described in Cattell (1946). Table 1 shows these variables and the reader will observe at each of the bipolar scale a collection or 'package' of descriptors accompanying the main descriptors. Each subject was rated as being either above or below average on each scale; product-moment correlations were then obtained and the correlation matrix was then factored using Thurstone's centroid method (e.g. see Cattell, 1952, pp. 35–60). We have not attempted to refactor the data; rather, our focus is on exploratory rerotation and our search begins with the unrotated factor matrix presented in full in the article.

As Cattell's analysis did not extend beyond nine (or, at the most, ten) identifiable factors, and as, in any case, the eigenvalues (see footnote on page 215) did not indicate even as many as this, the writer decided to commence the exploration on the generous side, principally in order to avoid any criticisms on the grounds of initial underfactoring (i.e. rotating too few factors). A full series of solutions was obtained for combinations of P and W which subsumed standard solutions, such as independent cluster, Varimax, and so on, and which ranged from orthogonal to quite markedly oblique solutions. Many hours were then spent comparing these solutions with each other, and with the original solution of Cattell. In addition, variance curves were plotted, and when inspected revealed support for an emerging opinion that only six or seven factors were potentially viable, as the variance always dropped considerably after this number of factors. Superior solutions were found in the partial oblique region recommended by Hakstian (1971) – by 'superior' is meant simply those solutions with the most factors having four or more acceptable salients, irrespective of any meaning subsequently attributed to the group of variables represented on any factor. But the impression was formed, after the expense of considerable time and energy on these ten factor rotations, that all the solutions represented overfactoring, i.e. rotating too many factors to simple structure. This is a point of divergence from Cattell who is concerned to include many tail factors of small variance in the rotation process claiming that greater clarity is thereby obtained. In respect of his view we began with ten-factor rotations but the (analytic) findings argued that one should rotate with fewer factors.

It was at about this time that an article by Digman (1972) – representing the fullest survey of the problem to that time – appeared suggesting that seven factors might be the correct number in rating data. Digman's work not only included a somewhat larger study than Cattell's (but on children where the factors might be fewer than in adults) but he had also analysed some of Cattell's rating studies on children, using his own methods which included use of the Harris–Kaiser method. However, as we were still under the influence of the Tupes–Christal–Norman suggestion, we initially explored a more restricted set of factors than that suggested by Digman's review.

Five- and seven-factor solutions

Following the protracted series of ten factor rotations, one of *five factor* rotations was undertaken because of (1) the literature on 'the recurrent five-factor structure (in ratings)' already referred to; (2) the impression gained, at this stage, that there were probably about five factors in the data (again see eigenvalues in footnote on page 215). One should mention that this 'five-factor structure' has been generated latterly from 20 and sometimes as few as 16 scales. It was hoped, however, that greater clarity would emerge from the *five-factor* rotations using the Hakstian H-K program by once more obtaining various standard solutions plus a number of partial and fully oblique solutions. In the event, and to the considerable surprise of the investigator, these restricted solutions were not better than the extended ones, and in both cases rather muddled anxiety and sociability factors emerged of the kind which Cattell might characterize as 'pseudo-second order' factors. This finding brings us to the point, and this should be quite forcefully presented, that analytic programs are little better than the decision as to the 'number of factors' – what, therefore, is the unskilled person, possibly using factor analysis for the first time (for, say, a thesis analysis) to do in view of the fact that there are more than a dozen alternative methods in the literature? The answer suggested here is that of empirical exploration, using as many clues – from the data, from the eigenvalues, from the literature, from projections and so forth – as possible. In the present instance, at this point, a view had emerged that *while ten was too many, five was too few*. Naturally, a full examination would then have taken us through all solutions in the range from five to ten; however, life was too short for this desirable option to be pursued so that we then aimed, deliberately, toward the eigenvalue-indicated region of seven factors (perhaps a little too high in the present instance, but paying respect to Digman's work). Apart from the necessity, which we had fulfilled, of paying operational respect to Cattell's views on the one hand, and to the more restricted set of Norman (1969) considerable time and effort would have been saved had the W parameter been appropriately set in the first instance.

It was at this stage of the investigation that an extraordinary contrast with the previous solutions (five and ten factor) emerged, for each of the seven-factor solutions which we now obtained, whether orthogonal or oblique, and excepting only the independent cluster solution, was markedly clearer than the previous ones, indeed in terms of salient aggregations (i.e. factor meaning) there was a family resemblance, as though the factors were hardening and becoming more resistant to the effects of this or that rotation. In terms of numbers of acceptable salients (see the writer's criterion above), the best of these solutions was of the kind advocated by Hakstian, i.e. a partial oblique solution with, in this case, the W parameter set at 3·5.

While one would like to present all these solutions in full, this is clearly an impractical proposition; therefore, the particular ('best') solution has been abstracted from the computer printout, as representative, for descriptive purposes (Table 1), with parameters $P = 0{\cdot}5$, $W = 3{\cdot}5$.

Table 1. *Factor salients* (*ratings*)

Variable number	Factor 1: Cooperativeness–Considerateness			Loading
24	*Demanding, Impatient* Expects a lot from other people. Makes constant demands. Self-centred. Forgets needs of others. Impatient.	vs.	*Emotionally mature* Self-sufficient without fuss. Thinks of others.	77
4	*Assertive, Self-assured* Assumes he can impose his (or her) will on others. Tends to lead or influence his associates. Tends to dominate. Tends to be boastful and assertive. Not held back by doubts. Invulnerable self-esteem.	vs.	*Submissive* Tends to let other people have their way. Tends to back down in a conflict. Humble, quiet, retiring. Not sure he is right. 'Embarrassable.'	69
12 (5)	*Good-natured, Easygoing* Does not mind when people use his property, time or energy. Generous, gives people 'the benefit of the doubt' when their motives are in question. Warm-hearted.	vs.	*Spiteful, Grasping, Critical* Gets irritable, 'awkward', or resentful if property or other rights are trespassed on. Inclined to be 'close' and grasping. Is generally surly, hard and spiteful.	−66
22 (6)	*Prone to jealousy* Becomes readily jealous of people. Unreasonably hostile.	vs.	*Not prone to jealousy*	63
19	*Hard, Stern* Toughly 'realistic' about problems. In extreme instances cynical. Looks at questions in a cold, objective fashion. Unaffected by personal and emotional appeals.	vs.	*Kindly, Soft-hearted* Tender-hearted. Cannot bear to see suffering unalleviated. In extreme instances sentimental and unrealistic. Of a grateful, understanding disposition.	62
11	*Suspicious* Believes rather too quickly that he is being unfairly treated. Imagines on insufficient grounds that certain people strongly dislike him. Inclined to brood over his troubles. Interprets things as having reference to himself when none is intended. Feels persecuted.	vs.	*Trustful* Free from suspicion.	57
15 (7)	*Mild, Self-effacing* Gentle-tempered. Blames himself (or nobody) if things go wrong.	vs.	*Self-willed, Egotistic* Goes his own way regardless of others. Blames others, not himself, whenever there is conflict or things go wrong. Headstrong. Predatory – tends to use other people for own ends.	−55
23	*Rigid* Always does things in one particular way. Is nonplussed if his routine of life is upset. Sticks to his own ideas and does not adapt to ways of doing things different from his own. Follows the letter of the law, not merely the spirit. Does not change and broaden with circumstances. Conservative in personal habit.	vs.	*Adaptable* Adapts his habits and ways of thinking to those of the group. Is satisfied with compromises – does not stand meticulously on the letter of the law. Is not upset, surprised, baffled, irritable if things are different from what he expected.	55

Table 1 (*cont.*)

Factor 1: Cooperativeness–Considerateness

Variable number				Loading
3 -	*Attention-getting* Shows off in company. Not happy unless in centre of the stage. Talks about self, accomplishments, important friends, etc. Likely to show some 'affected' behaviour.	vs.	*Self-sufficient* Not under compulsion to impress or to get sympathy or attention.	44

Variable number in Norman's study is shown in parentheses. Norman had 16 variables, four markers per factor. For example, if our first factor resembles Norman's second, some of his markers, 5, 6, 7, 8, should appear above. This factor, therefore, resembles Norman's *Agreeableness* factor.

Factor 2: Phlegmatic vs. Surgent

9	*Languid, Slow* Lacks vigour. Has an absent-minded manner. Vague and slow in speech, etc.	vs.	*Energetic, Alert* Vigorous, quick, and alert most of the time. Spirited. Enterprising. Always 'on the ball'.	−75
16 (1)	*Silent, Introspective* Says very little; gives the impression of being introspective and occupied with thoughts.	vs.	*Talkative* Talks a lot, to everybody.	−71
5	*Depressed, Solemn* Earnest and solemn most of the time. Not easily moved to laughter. Seeming slow and depressed rather frequently.	vs.	*Cheerful* Generally bubbling over with good cheer. Optimistic. Enthusiastic. Prone to cheerful, witty remarks. 'Laughterful.'	−65
13	*Calm, Phlegmatic* Remains relatively calm in dispute, danger, social hilarity, temptation. Shows few signs of emotional excitement. (Not merely controls impulses, but scarcely seems to have them.)	vs.	*Emotional* Gets emotional (anger, fear, jollity, sex, sorrow or disgust) on slight provocation. Frequently emotional and passionately excited. Shows marked signs of emotion, even if controlled.	−56
18 (3)	*Cautious, Retiring, Timid* Avoids the strange and new. Looks at all aspects of a situation overcautiously. Keeps clear of difficulties. Uninquiring, lacking in desire to try new things.	vs.	*Adventurous, Bold* Rushes in carefree fashion into new experiences, situations, emergencies. Ascendant: ready to meet anything. Happy-go-lucky. Has a great appetite for life.	−50
15	*Mild, Self-effacing*	vs.	*Self-willed, Egotistic*	−42

The numbers 1 and 3 given in parentheses refer to the variable numbers among the 16 scales used by Norman (1969) and represent two of the markers 1, 2, 3, 4 for his Extraversion–Surgency factor.

Factor 3: Emotional maturity

35	*Dependent, Immature* Emotionally and intellectually dependent on others. Generally adopts the opinion of the group or of authority without much thought. Thoughts vague and confused. Rather immature.	vs.	*Independent-minded* Thinks things out for himself and adopts a clear and definite independent position. Tends to be a leader in discussion. Is interested in public opinion and how to shape it.	−74

Table 1 (*cont.*)

Variable number				Loading
	Factor 3: Emotional maturity			
10	*Boorish* Rather ignorant. Unreflective. Does not read much or enjoy intellectual problems. Narrow, simple interests.	vs.	*Intellectual, Cultured* Has wide interest and knowledge, especially in intellectual matters. Is thoughtful and introspective about life. Enjoys analytical, penetrating discussions in small groups.	−67
21	*Polished* Polite and charming in social situations. Deals with people gracefully and skilfully. Refined with speech, manner, etc. Familiar with good etiquette.	vs.	*Clumsy, Awkward* Clumsy in social situations. Crude in speech, manner, etc.	64
2	*Emotionally stable* Can be depended upon to look at questions objectively, without emotional prejudice, and in the same constant light from day to day. Above emotion in his judgements. Dependable and realistic.	vs.	*Changeable* Sees things in terms of the emotion of the moment. Emotional bias changes from day to day and place to place. Does not remain the same person from day to day. Undependable.	57
28	*Composed* Quite free from shyness. Completely at home in company or with strangers.	vs.	*Shy, Bashful* Not necessarily unsociable, but extremely shy and sensitive. Goes out of way to avoid meeting people. Tense when speaking in a group. Tries to avoid occasions when he will be in limelight.	46
17 (12)	*Persevering, Determined* Sees a job through in spite of difficulties or temptations. Strong-willed. Persisting in his motives. Painstaking and thorough.	vs.	*Quitting, Fickle* Gives up rather easily. Led astray from main purposes by stray impulses. Slipshod – does not finish a job thoroughly.	45
	Factor 4: Extraversion (sociability)			
34 (4)	*Gregarious, Sociable* Likes to be in large groups. Seeks people out for the sake of company. Likes parties as often as possible. Not fond of being alone.	vs.	*Self-contained* Does not seem to miss company of others. Goes own way.	68
33 (2)	*Frank, Expressive* Comes out readily with his real feelings on various questions, so that you know where you stand with him. Expresses his feelings, sad or gay, easily and constantly. Easy to understand.	vs.	*Secretive, Reserved* Keeps his thoughts and feelings to himself. Often leaves you puzzled as to the motives of his actions. Inscrutable. Does not give away information for the fun of it.	64
32	*Marked interest in opposite sex* Dates a good deal. Shows much interest in opposite sex.	vs.	*Slight interest in opposite sex* Dates very little (and not merely from lack of opportunity).	59

Table 1 (*cont.*)

Variable number	Factor 4: Extraversion (sociability)			Loading
7	*Attentive to people* Interested in people, their troubles, their personalities. Makes friends with people and remembers their personal interests. Spends much time in dealing with people.	vs.	*Cool, Aloof* Tends to be indifferent to, and to ignore, people. Gives the impression of brooding on his own thoughts or of being cold and indifferent.	50
28	*Composed*	vs.	*Shy, Bashful*	43

Two of Norman's Extraversion–Surgency markers are indicated by the numbers 2 and 4 in parentheses.

Factor 5: Superego

6 (10)	*Frivolous* Does not seem to take responsibilities seriously. Undependable. Thoughtless. Refuses to accept responsibilities of his age.	vs.	*Responsible* Has a sense of responsibility to his parents, community, etc. Can be depended upon to be loyal to agreed standards. Trustworthy.	−70
25	*Unconventional, Eccentric* Breaks conventional rules. Does not mind being different from other people in dress, manners, interests. Has well-marked individual ways (of which he is aware) and even fads and eccentricities.	vs.	*Conventional* Conforms to the accepted standards and ways. Seems distressed if he finds he is being 'different'.	−53
27 (11)	*Conscientious* Careful about principles of conduct. Guided by ideas of truthfulness, honesty, unselfishness. Scrupulously upright where personal desires conflict with principle.	vs.	*Somewhat unscrupulous* Inclined to somewhat shady transactions. Not too careful about right and wrong where own wishes are concerned. Not particularly just, honest or unselfish.	51
17 (12)	*Persevering, Determined*	vs.	*Quitting, Fickle*	48
20 (9)	*Insistently orderly* Tidy, over-precise, especially over details. Drives other people to be the same. Strict, fussy, pedantic. Insists on everything being orderly. (In these respects rather 'uncomfortable to live with'.) Seems unable to relax. Miserly.	vs.	*Relaxed, Indolent* Rather careless of detail. Lazy. Careless over expenditures. Has no difficulty in relaxing. Enjoys ease.	40

Variables also in Norman's *Conscientiousness* factor are shown by the numbers 9, 10, 11 and 12 in brackets.

Factor 6. Adjustment–Emotionality

26 (14)	*Placid* Calm, tough. 'What's the fuss about?' attitude.	vs.	*Worrying, Anxious* Worries constantly, sensitive, hurried; seems to suffer from anxieties without adequate cause; slight suppressed agitation much of the time.	−70

Table 1 (*cont.*)

Variable number	Factor 6. Adjustment–Emotionality			Loading
30 (13)	*Neurotic fatigue* Seems to get tired and overwrought. Is irrationally irritable. Jumps when spoken to. Shows facial tics and other signs of 'nervousness' (e.g. fidgeting, tremor, digestive disturbances, poor memory). Constantly complains of fatigue.	vs.	*Absence of neurotic fatigue*	68
8 (15)	*Easily upset* Easily embarrassed or put off balance in conversation. Gets confused in emergency. Blushes, shows excitability, becomes incoherent. (Not general emotionality, but momentary 'nervousness').	vs.	*Unshakeable poise, Tough* Self-possessed, hard, does not lose composure, e.g. through emotional provocation.	64
14 (16)	*Hypochondriacal* Dwells on illness or hurts a good deal. Is afraid of going to die when has relatively trivial illness. Fusses a good deal over bodily symptoms.	vs.	*Not so* Absence of H.	53

(This factor is identical with Norman's *Emotional stability* factor, in that all four of his markers are here as variables 13, 14, 15, 16 in parentheses.)

FACTOR DESCRIPTION

Let us now turn to the question of what factors (and how many) are to be found in Cattell's personality sphere study, using analytic rotation of the kind described.

Cooperative–Considerateness (*Agreeableness*)

The factor of largest variance generally resembled Norman's (1969) Agreeableness factor, although here the two leading salients indicated a possible alternative title which accords with that previously assigned by the writer to a well-replicated questionnaire factor, and is so named to be consistent with previous articles and with the test manual.* This first factor was much easier to identify *than any* of those in Cattell's original solution because it had nine large salients. It accounted for 12·3 per cent of the variance.

Phlegmatic vs. Surgent

The six salients for this factor included two out of the four Norman variables which comprised his factor Extraversion–Surgency. The numbers in Cattell's matrix are again shown to enable the reader to make appropriate comparisons in the literature. Because of the presence of variables 5, 9, 13 (Norman's variable numbers are shown in brackets) which indicate descriptors at one pole, languid, slow, solemn, depressed, we have adopted the present title, and the factor, be it noted, resembles the 'Surgency vs. Repression' factor in Harman & French (1973). It accounted for 10·1 per cent of the variance.

---

* Technical background and user information for the Howarth Personality Questionnaire; Individuality Inventory; Adjunct Personality Factor Inventory; Howarth Mood Adjective Check List (HPQ/INDINV/APF/HMACL). Replacing: Howarth, Edgar *Manual for the H.P.Q.* Personality Research Associates, 1973.

Emotional maturity

That factors can emerge not only in terms of variables included in the study, but also in terms of sample characteristics, hardly requires drawing to the reader's attention. In the present case the group tested consisted of youngish students, and contrasts with the more mature subjects of the original Harvard study. With greater possibility of variation in emotional maturity, it is not surprising, therefore, that this factor is quite prominent in the solution. While neither Cattell nor Norman reported such a factor, one of the same title was described by Harman & French (1973) which included these three aspects:

(1) Patient, adjusts to frustration vs. verbally aggressive, demanding.

(2) Modest, shuns attention vs. self-centred, egotistical.

(3) Satisfied, cooperates with authority vs. asserts independence from authority, stubborn.

Their description of the factor contents appears to conform more with our first factor, in some respects, and in the present case we found personality descriptors perhaps more fitted to our title such as immature, boorish, clumsy, quitting, changeable versus their opposites for each scale. This factor accounted for 9·7 per cent of the variance.

Extraversion (Sociability)

As previously mentioned, this fourth factor correlated somewhat with the second factor, and just as two of Norman's markers were found in the latter, so the remaining two (he had four in all) markers appeared on this fourth factor. Clearly, both in terms of this splitting of Norman's markers, and in view of the correlation observed, one might choose to regard factors two and four jointly as 'extraversion'; however, the second factor might be held to represent the surgency aspect and the present factor the sociability aspect of 'extraversion'. One might then speculate that (1) the surgency factor might best represent a basic temperamental trait; (2) the sociability factor could be considered a part of a broader factor called 'gregariousness' by Harman & French (1973) and possessing three aspects. This factor accounted for 8·6 per cent of the variance.

Superego (Conscientiousness)

This factor resembles both the ego-ideal factor of Adcock & Adcock (1967) and the conscience factor reported by Howarth & Browne (1971*a*). It also includes among the leading five salients three of four of Norman's variables for his 'Conscientiousness' factor. This factor accounted for 8·3 per cent of the variance.

Emotional stability

The salient aggregation found appears to represent emotional control or stability, and resembles rubrics such as 'neuroticism' or 'anxiety'. The reason for not simply using either of the latter terms is that either one is by no means 'simple' as evidenced by the more than 3000 literature references variously attempting to define these terms. What we found was an almost perfect match for Norman's 'emotional stability' factor, as all four of his markers were herein recovered. Moreover, the present factor resembles the 'emotional stability' factor of Harman & French (1973). This factor accounted for 7·9 per cent of the variance.

Unidentified

As seven factors take us somewhat beyond the 'unit eigenvalues' rule (i.e. the Kaiser–Guttman criterion) and as there were only three loadings greater than 0·40, this factor remains unidentified. As a nice technical point, one prefers to go a *little* beyond the Kaiser–Guttman (Kaiser, 1961) in order to achieve greatest resolution in the leading factors.

DISCUSSION

We have now adduced sufficient evidence to address ourselves to the question as to whether Cattell genuinely mapped the personality sphere in the first instance, and the connected issues as to whether he was justified in (*a*) further extending the basic system; (*b*) continuing for such a long period to expound and defend that system. This evidence would appear to indicate that Cattell overfactored in his initial work, as by using objective methods we have found considerable agreement with sources outside his system. While remaining in considerable sympathy with Cattell's strategy *vis-à-vis* (*a*) multivariate definition and (*b*) comprehensive choice of variables (still, be it noted, 30 years later, there have been very few studies of a more comprehensive kind), it must be said that there are questionable aspects to his tactics which he is still entitled to defend, but which this writer and others are not willing to follow or support. Indeed, rather than system defence, what is now needed is a movement of investigators toward a consensus, buttressed by coordinated studies and the recent survey of past work by Harman & French (1973) should offer a guide for this. As a positive suggestion, and because of improved methods for large-scale factor analysis, much more comprehensive studies should be carried out. When these are done it is to be expected that there will be less grounds of, or possibility for, disagreements among investigators (Howarth, 1972).

If Cattell did not accurately pinpoint personality factors it behoves us to ask two questions: (1) Has anyone else satisfactorily resolved the issue of how many and what kind of personality factors are there? (2) By what means could this be accomplished? Perhaps the closest to date is the consensus by Harman & French (1973) but even therein many of the studies make no claim to be of a comprehensive nature. In regard to the second question, the following steps should be followed:

(1) The (rating) scales should be marked by adjectival descriptors from current sources containing contemporary data on meaningfulness, frequency and social desirability.

(2) Bipolar scales should be carefully pre-tested in order to determine both clarity of meaning and semantic opposition.

(3) Frequency of use indices should be applied; this is of special importance where less educated raters are to be used.

(4) Elimination of social desirability extremes.

(5) Testing of populations from various social levels.

REFERENCES

ADCOCK, C. J. (1972). Multivariate measurement theory. In R. M. Dreger (ed.), *Multivariate Personality Research*. New Orleans: Claitor Press.

ADCOCK, N. V. (1974). Testing the test: How adequate is the 16 PF with a New Zealand student sample? *New Zealand Psychologist* **3**, 11–15.

ADCOCK, N. V. & ADCOCK, C. J. (1967). A factorial study of the ego reference system. *J. genet. Psychol.* **110**, 105–115.

ADCOCK, N. V., ADCOCK, C. J. & WALKEY, F. H. (1974). Basic dimensions of personality. *Int. Rev. appl. Psychol.* **23**, 131–137.

BERG, I. A. (1972). Raymond B. Cattell. The man, his explorations and his impact on psychology. In R. M. Dreger (ed.), *Multivariate Personality Research*. New Orleans: Claitor Press.

CATTELL, R. B. (1946). *The Description and Measurement of Personality*. New York: World Press.

CATTELL, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika* **12**, 197–220.

CATTELL, R. B. (1952). *Factor Analysis*. New York: Harper.

CATTELL, R. B. (1973). *Personality and Mood by Questionnaire*. San Francisco: Jossey-Bass.

CATTELL, R. B., EBER, H. M. & TATSUOKA, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, Ill.: I.P.A.T.

DIELMAN, T. E., CATTELL, R. B. & WAGNER, A. (1972). Evidence on the simple structure and factor invariance achieved by five rotational methods on four types of data. *Multivariate behav. Res.* **7**, 233–242.

DIGMAN, J. F. (1972). The structure of child personality as seen in behaviour ratings. In R. M. Dreger (ed.), *Multivariate Personality Research*. New Orleans: Claitor Press.

EYSENCK, H. J. (1972). Primaries or second order factors: A critical consideration of Cattell's 16 PF battery. *Br. J. soc. clin. Psychol.* **11**, 265–269.

EYSENCK, H. J. & EYSENCK, S. B. G. (1969). *Personality Structure and Measurement*. London: Routledge & Kegan Paul.

FISKE, D. W. (1971). *Measuring the Concepts of Personality*. Chicago: Aldine.

GOLDBERG, L. R., NORMAN, W. T. & SWARTZ, E. (1972). The comparative validity of questionnaire data (16 PF scales) and objective test data (O–A battery) in predicting five peer-rating criteria. *Oregon Research Institute Res. Bull.* **12**, 1–18.

GREIF, S. (1970). Untersuchungen zur deutschen Übersetzung des 16 PF Fragebogens. *Psychol. Beitrage* **12**, 186–213.

HAKSTIAN, R. (1970). A computer program for oblique factor transformation using the generalized Harris–Kaiser procedure. *Educ. psychol. Measurement* **30**, 703–705.

HAKSTIAN, R. (1971). A comparative evaluation of several prominent methods of oblique factor transformation. *Psychometrika* **36**, 175–193.

HARMAN, H. H. & FRENCH, J. W. (1973). *Toward the Establishment of Noncognitive Factors through Literature Search and Examination*. New Jersey: Educational and Testing Service (ETS).

HARRIS, C. W. & KAISER, H. F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika* **29**, 347–362.

HOWARTH, E. (1972). Factor analysis has only just begun to fight – a reply to Lykken. *J. exp. Res. Person.* **6**, 268–272.

HOWARTH, E. & BROWNE, J. A. (1971*a*). An item factor analysis of the 16 PF. *Personality: An International Journal* **2**, 117–139.

HOWARTH, E. & BROWNE, J. A. (1971*b*). Investigation of personality factors in a Canadian context. *Can. J. behav. Sci.* **3**, 161–173.

HOWARTH, E., BROWNE, J. A. & MARCEAU, R. (1972). An item analysis of Cattell's 16 PF. *Can. J. behav. Sci.* **4**, 85–90.

HOWARTH, E. & CATTELL, R. B. (1973). The multivariate contribution to personality research. In B. B. Wolman (ed.), *Handbook of General Psychology*. New York: Prentice-Hall.

KAISER, H. F. (1961). A note on Guttman's lower bound for the number of common factors. *Br. J. statist. Psychol.* **14**, 1–2.

KARSON, S. & O'DELL, J. W. (1974). Is the 16 PF factorially valid? *J. Person. Assessment* **38**, 104–114.

LEVONIAN, E. (1961). A statistical analysis of the 16 personality factor questionnaire. *Educ. psychol. Measurement* **21**, 589–596.

NORMAN, W. T. (1969). To see ourselves as others see us: Relations among self-perceptions, peer perceptions, and expected peer-perceptions of personality attributes. *Multivariate behav. Res.* **4**, 417–443.

SELLS, S. B., DEMAREE, R. G. & WILL, D. P., Jr. (1970). Dimensions of personality. I. Conjoint factor structure of Guilford and Cattell trait markers. *Multivariate behav. Res.* **5**, 391–422.

TUPES, E. C. & CHRISTAL, R. E. (1961). Recurrent personality factors based on trait ratings. USAF Lackland Air Force Base, Personnel Laboratory Technical Report ASD-TR-61-97.

(*Manuscript received* 7 *May* 1974; *revised manuscript received* 20 *June* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 231–242
Printed in Great Britain

HOW REPRESENTATIVE ARE CONCEPT ATTAINMENT EXPERIMENTS?*

BY N. E. WETHERICK
University of Aberdeen

AND R. L. DOMINOWSKI
University of Illinois at Chicago Circle

Two experiments are reported which differ from orthodox concept attainment experiments in employing tasks in which the number of attributes exemplified by each instance is clearly larger than the number about which information is available and in which there is no guarantee that a solution will be found in terms of the attributes about which information was presented initially. The results suggest that the orthodox task fails in important respects to represent the corresponding real-world activity.

Experiments in concept attainment typically involve the presentation to the subject of instances exemplifying a small number of attributes, each attribute taking one of a small number of values. Bruner, Goodnow & Austin (1956) employed reception tasks (in which a sequence of instances is presented, chosen by the experimenter) and selection tasks (in which the instances are chosen by the subject) and one or other of these two basic procedures have usually been employed in more recent work in the area. Such experiments seek to represent the real-world activity in which instances of a particular kind are known to fall into one of two categories, positive (having a particular characteristic) and negative (not having that characteristic), and the subject tries to find a simple perceptual attribute or set of attributes, possessed by all the positive instances but not by any negative instance, that will enable him to predict which instances will turn out positive.

Experimental tasks and the real-world activities they seek to represent necessarily differ. This will be true of any task by comparison with its corresponding segment of reality, but the possibility must be borne in mind that a task may be unrepresentative in some crucial respect. If this is the case it can generate no useful advance in our knowledge. The experiments reported here try to establish whether some fairly obvious differences between the tasks usually employed in the experimental investigation of concept attainment and the corresponding real-world activity are crucial in this sense.

The first difference to be considered follows from the fact that the subject can take in information about only a limited number of attributes of each instance. In an experimental task it will be obvious which attributes are to be considered and there are usually few enough of them to ensure that the subject's perceptual capacity is not exceeded. In real life, however, the number of attributes exemplified by an instance will be indefinitely large. Since the subject can only take in information about a limited number of them he will have to select a set of attributes to be observed on the basis of past experience, or the salience of particular attributes, or at

---

* A preliminary report of the experiments was given at the annual conference of the British Psychological Society (Bangor, 1974).

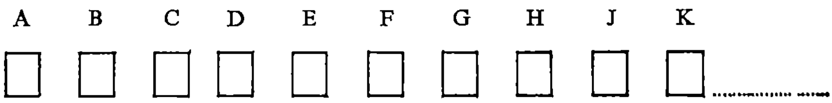


Fig. 1. Instances as presented, either one at a time on cards or cumulatively on a sheet of paper. In each box (representing an attribute) the number 1, 2 or 3 appeared. The words 'positive' or 'negative' were entered on the dotted line to the right.

random. What will be the effect of drawing the subject's attention (as in real life) to the existence of attributes which are not being observed and can only be observed by dropping from consideration some attributes that are being observed?

The second difference is a consequence of the first. In the experimental task it is given, implicitly or explicitly, that the predictor attribute will be one of those to which the subject's attention has been drawn. In real life there may be no predictor attribute and there will certainly be no guarantee that any selected set of attributes to be observed will include it.

The third difference lies in the fact that, in the attainment of concepts, information derived from previous instances has to be carried in the memory. In an experimental task it is of course possible to withdraw each instance after it has been presented or allow the subject to keep the series in view throughout but the differences already considered combine to make the problem of memory load more serious in real life than it can be in the laboratory.

The fourth and last difference follows from the fact that, while very often the values of attributes exemplified by an instance are known before it becomes apparent whether the instance is positive or negative (this is normally the case in experimental science), it may be necessary or advisable outside the laboratory to select instances because they are positive or negative and then look to see what values they exemplify of particular attributes. What difference (if any) does this make?

In the experiments instances consisted of rows of ten boxes (representing attributes) as illustrated in Fig. 1. In each box the number 1, 2 or 3 appeared. The word 'positive' or 'negative' was entered on the dotted line to the right. A series of 24 instances was generated in which one attribute value (H3) predicted membership of the class of positive instances. The instances were presented in random order but there were an equal number of positives and negatives. The first instance was positive and no run of more than two positives or negatives occurred. All the positive instances had number 3 in box H by definition; in the negative instances 1 and 2 appeared in box H with equal probability. In the remaining nine boxes, the numbers 1, 2 and 3 appeared randomly with equal probability. An attribute could be eliminated as a predictor of membership of the class of positive instances if the same number appeared in the box in a positive and a negative instance or if different numbers appeared in two positive instances. In this way Instance 2 (negative) eliminated two attributes and Instance 3 (positive) eliminated five, one further attribute was eliminated by Instance 5 (negative). Up to and including Instance 11, E2 and H3 were equally tenable hypotheses but E2 was eliminated by Instance 12 (positive). The instances were presented either one at a time on cards (placing maximum load on the subject's memory) or cumulatively on a sheet of paper (placing minimum load on memory). Information could be given about different numbers of attributes of each instance; determined either by the experimenter or by the subject (who could indicate attributes about which he desired information by underlining the appropriate box). After an instance had been presented the subject was asked to state what hypotheses he was considering and how strongly he held each hypothesis using a scale ranging from 1 (hypothesis 'under consideration') to 5 (subject 'certain' that the hypothesis is correct).

EXPERIMENT I

In Expt. I the subject was given information about four attributes of each instance. In instances subsequent to the first he selected the four attributes for himself but in the first instance they were selected by the experimenter, the subject being told that the selection had been random and that consequently the set of four might or might not include the single predictor attribute he was looking for.

The experiment was concerned with two particular aspects of concept attainment. Half the subjects were given ADFJ as their initial attribute set, half were given ADFH (including 'H', the predictor). It was therefore possible to detect any effect on performance of including the predictor in the initial set. Half the subjects were told whether each instance was positive or negative before being asked to select their attributes. The other half were given this information after they had been given the values of the attributes selected. It was therefore possible to detect any effect on attributes selected of knowing in advance that the instance was positive or negative.

Subjects

Thirty-two undergraduate students not taking psychology, aged 18–23 years (16 male, 16 female). Subjects were paid for participation.

Procedure

Instances were presented cumulatively (placing minimum load on memory). It was emphasized that one attribute value (i.e. one number in one of the boxes) predicted membership of the class of positive instances and that the set of attributes of the initial instance might or might not include the predictor. Subjects were asked to state their hypotheses after each instance, together with the strength with which they held each hypothesis and were told that if they proposed a hypothesis with strength 5 (indicating that they were 'certain' that the hypothesis was correct) they would be told whether or not it was correct and if it was they would have completed the task. They were, however, asked not to propose hypotheses with strength 5 unless they were absolutely 'certain'.

Subjects were introduced to the type of problem with the aid of a training problem in which instances consisted of rows of four boxes only and information was given about all four. In the training problem Instance 3 (negative) eliminated one attribute, Instance 4 (positive) eliminated another and Instance 8 (positive) eliminated a third, leaving B1 as the predictor.

Results

It was logically possible to solve the training problem after Instance 8 had been presented. Eighteen of the 32 subjects solved it at that point. The instructions did not preclude guessing (i.e. proposing a hypothesis with strength 5 before competing hypotheses had been eliminated), and 10 subjects guessed the solution with fewer than eight instances. The remaining four subjects required 9 (2), 12 and 20 instances respectively. Care was taken to ensure that all subjects understood the procedure, irrespective of their performance.

The main problem was more difficult in that information was given not about all four attributes of an instance but about four out of ten attributes. The problem could not logically be solved till Instance 12 had been presented and Table 1 shows that

Table 1. *Performance of four groups of subjects* ($n = 8$ *in each group*) *on a concept task* (*Expt. I*)

	Succeed with				Mean number of attributes un-eliminated
	Fewer than 12 instances	12 instances	More than 12 instances	Fail (24 instances)	
Group 1 Predictor not included in initial set. Next instance not known to be positive/negative	1	2	3	2	0·50
Group 2 Predictor not included in initial set. Next instance known to be positive/negative	1	5	1	1	0·13
Group 3 Predictor included in initial set. Next instance not known to be positive/negative	3	3	2	0	2·00
Group 4 Predictor included in initial set. Next instance known to be positive/negative	5	1	2	0	2·88
Total	10	11	8	3	—

11 subjects solved it at that point, 10 guessed on the basis of fewer than 12 instances, 8 required more than 12 instances and 3 had not succeeded after 24 instances.

It appears from a comparison between the results of Groups 1 and 2 (predictor not included in the initial set) and Groups 3 and 4 (predictor included) that the effect of including the predictor was significantly to increase both the incidence of 'guessed' solutions ($\chi^2 = 5{\cdot}24$, $P < 0{\cdot}025$) and the mean number of attributes that had not been eliminated when the solution was proposed ($\chi^2 = 29{\cdot}03$, $P < 0{\cdot}001$). Presumably the subject will begin by investigating the possibility that the predictor is one of the attributes in his initial set. The question is how many instances consistent with a hypothesis should be observed before proposing it as the solution (bearing in mind that no finite number *proves* the hypothesis) and how careful need one be to ensure that competing hypotheses have been eliminated (bearing in mind that in real life all possible competing hypotheses will not even be under consideration). The way in which this dilemma was resolved seems to have depended on the nature of the initial attribute set. If one considers hypotheses advanced up to and including Instance 11 (i.e. when both H3 and E2 were tenable hypotheses) then 14 out of 16 subjects in the 'predictor included' groups (3 and 4) considered H, as against only 9 out of 16 in the 'predictor not included' groups (1 and 2). This difference is just significant ($\chi^2 = 3{\cdot}86$, $P < 0{\cdot}05$). Six as against 8 considered E and 2 as against 6 considered some other attribute. E was not favoured by being included in the initial attribute set and was considered more or less equally by all four groups. H was considered significantly more frequently by the groups in whose initial set it was included but tended to be 'guessed' as the solution before competing hypotheses had been eliminated.

The second aspect of the situation with which this experiment was concerned was

Table 2. *Mean number of attributes chosen, common to instances* 1–2, 1–3, *etc., when the next instance is known or not known in advance to be positive or negative* (*Expt. I*)

	Attributes common to instances:			
	1–2	1–3	2–3	1–2–3
Next instance known to be positive/negative (Groups 2 and 4)	2·63	2·31	2·50	1·75
Next instance not known to be positive/negative (Groups 1 and 3)	2·25	2·25	1·56	1·25

the effect on choice of attributes of knowing in advance that the next instance would be positive or negative. Half the subjects, (Groups 2 and 4) were given this information and half (Groups 1 and 3) not. Attributes chosen for Instance 2 and Instance 3 were examined in detail. Table 2 shows the mean number chosen common to Instances 1 and 2, 1 and 3, 2 and 3, and 1, 2 and 3. This mean number amounts to an index of conservatism, i.e. relative unwillingness to move away from the attributes of the preceding instance. Subjects who knew in advance appear to have been significantly more conservative throughout (Kolmogorov–Smirnov, $D_{\max} = 8$, $P < 0{\cdot}01$, two-tailed). The second instance was in fact negative and five subjects from Groups 1 and 3 selected four completely new attributes for this instance. No subject from Groups 2 and 4 behaved in this way. Comments from some subjects to the effect that 'negative instances are no help' suggest that these five may have been gambling that the second instance would be positive; on the assumption that, if it was, useful information would be maximized. In fact it is the comparison between pairs of instances that enables attributes to be eliminated and, for this purpose, pairs consisting of one positive and one negative instance are as useful as pairs consisting of two positives. In the present context conservatism pays and it is of interest that (if our interpretation is correct) drawing subjects' attention in advance to the fact that an instance is positive or negative reduces the tendency to rely on positive instances only. This parallels a finding in Wetherick (1966) where the same precaution sharply reduced incidence of a type of error in which sets of mixed positive and negative instances were treated as if they were all positive.

Group 2 was clearly the most efficient overall – in this group the predictor attribute was not included in the initial set (encouraging the subject to consider all possibilities) and it was known in advance whether the next instance would be positive or negative (encouraging conservatism in the choice of attributes). The result was a high level of performance and a low incidence of uneliminated attributes. In Group 1, although the predictor attribute was not included, the next instance was *not* known in advance to be positive or negative, the result was a lower level of performance but with a low incidence of uneliminated attributes. Groups 3 and 4 have a high level of performance but a high incidence of 'guessed' solutions and uneliminated attributes.

EXPERIMENT II

The second experiment employed the problems used in Expt. I with modifications. Terminating the experiment as soon as the correct hypothesis was proposed with strength 5 had proved wasteful of information since the subject sometimes guessed correctly early in the series. In

Expt. II the subject could still propose hypotheses with strength 5 but was told that an (unstated) number of instances would be presented in any case. The first 12 instances only were presented and subjects were told whether an instance was positive or negative at the same time as they received information about the attributes it exemplified.

Two further aspects of concept attainment were selected for investigation in Expt. II. The effects of memory load (i.e. the need to remember information derived from instances that had been presented earlier in the series) and the effects of size of attribute set (i.e. the number of attributes of each instance about which information was available).

Subjects

Thirty-six normal adults chosen from the Aberdeen Subject Panel, aged 18–35 years (15 male, 21 female). All the subjects had received some tertiary education and were paid for participation.

Procedure

Each subject attended on two occasions with an interval of two to three weeks between. On the first occasion he/she did the training problem and then the main problem; instances in both problems being presented one at a time on cards. Information obtained from previous instances had on this occasion to be held in the memory. The subject was given no information about his performance on the main problem. On the second occasion he/she was asked to do a problem 'similar to the one you did before but simpler, making smaller demands on memory' and was in fact presented cumulatively with the sequence of instances selected on the first occasion. For each subject it was therefore possible to determine what information could logically be derived from the instances selected, what information had in fact been derived under conditions of maximum memory load and what information had been derived under conditions of minimum memory load. Subjects were divided into four groups. In Group 1 information was made available about four attributes of each instance (as in Expt. I) and the initial instance was ADEJ. In Group 2 information was made available about seven attributes of each instance – initial instance ABDEGJK. In Group 3 information was made available about all ten attributes of the initial instance and thereafter about any number of attributes at the subject's own choice. In Group 4 information was made available about all ten attributes of every instance. In the first two groups E was included in the initial set but H was not; since, however, the hypotheses E2 and H3 were equally tenable until Instance 12 was presented (eliminating E2), the effects of including the predictor in the initial set could be evaluated by comparing the attention paid to E (included) and to H (not included) in Groups 1 and 2.

Results

Sixteen of the 36 subjects solved the training problem (i.e. proposed the correct hypothesis with strength 5) after presentation of Instance 8, 2 guessed the solution with fewer than 8 instances and 8 succeeded with between 9 and 12 instances. Ten subjects failed but most of them were considering the correct hypothesis although they had not yet proposed it with strength 5. As in Expt. I, care was taken to ensure that all subjects understood the procedure.

Table 3 shows the results obtained with the main problem. Each subject's responses were analysed to determine whether he could be regarded as having obtained a 'logically correct' solution, a 'correct' solution or an 'incorrect' solution. In order to be classified as 'logically correct' the subject had, after Instance 12, to be entertaining H3 as a likely hypothesis (though not necessarily with strength 5), to be entertaining any other hypothesis that had not been eliminated by his sequence of instances at some strength equal to or less than that of H3 and had not to be entertaining any hypothesis that had been eliminated by his sequence of instances. A subject who met the first of these criteria but not all three was classified as 'correct', other subjects were classified as 'incorrect'. The figures in parentheses in Table 3

Table 3. *Performance of four groups of subjects* ($n = 9$ *in each group*) *on a concept task* (*Expt. II*)

(Figures in parentheses indicate what the outcome would have been if E2 had been correct instead of H3.)

	Maximum memory load			Minimum memory load		
	Logically correct	Correct	Incorrect	Logically correct	Correct	Incorrect
Group 1 (4 atts. out of 10)	0 (0)	4 (5)	5 (4)	2 (2)	4 (4)	3 (3)
Group 2 (7 atts. out of 10)	3 (3)	4 (2)	2 (4)	5 (5)	3 (3)	1 (1)
Group 3 (Subject's own choice)	4 (4)	1 (2)	4 (3)	6 (4)	2 (2)	1 (3)
Group 4 (10 atts. out of 10)	2 (2)	3 (3)	4 (4)	5 (5)	3 (2)	1 (2)
Total	9 (9)	12 (12)	15 (15)	18 (16)	12 (11)	6 (9)

Table 4. *Number of valid and invalid hypotheses proposed* (*Expt. II*)

	Maximum memory load		Minimum memory load	
	Valid hypotheses	Invalid hypotheses	Valid hypotheses	Invalid hypotheses
Group 1 (4 atts. out of 10)	10·2	4·4	21·1	4·9
Group 2 (7 atts. out of 10)	15·1	7·1	27·0	7·2
Group 3 (Subject's own choice)	12·0	3·5	18·3	1·7
Group 4 (10 atts. out of 10)	13·5 (14·5)*	4·9 (4·0)*	22·2 (21·4)*	4·4 (0·9)*

* These figures exclude the results of one subject (S.32) who proposed an exceptionally large number of 'invalid' hypotheses.

show what the result would have been if E2 had been correct, i.e. if Instance 12 had supported E2 and eliminated H3 instead of vice versa. This involved a wholesale reshuffle of subjects between categories but the overall picture is, as would be expected, much the same. Subjects appear to have been able to make better use of information obtained under conditions of minimum memory load ($\chi^2 = 14{\cdot}40$, $P < 0{\cdot}001$ for H3 correct; $\chi^2 = 7{\cdot}92$, $P < 0{\cdot}005$ for E2 correct). At the end of their first session (maximum memory load) subjects were asked to classify as positive or negative 12 further instances, information being given about all ten attributes of each instance. All but two of the 21 subjects classified as 'logically correct' or 'correct' completed this additional task without error. One of these two classified the first ten correctly and then seemed to forget the criterion, the other was only technically correct since he seems to have guessed H on the last acquisition trial. This confirms that subjects who were 'logically correct' or 'correct' were capable of classifying on the basis of their hypothesis even though the acquisition task did not require them to do so.

Table 5. *Number of hypotheses proposed up to and including Instance* 11, *showing the amount of attention paid to 'H', 'E' and 'other' attributes* (*Expt. II*)

	Maximum memory load	Minimum memory load
Group 1 (4 atts. out of 10)		
'H'	15	35
'E'	41	53
'Other'	65	124
Group 2 (7 atts. out of 10)		
'H'	14	63
'E'	37	69
'Other'	111	156
Group 3 (subject's own choice)		
'H'	43	55
'E'	47	57
'Other'	41	57
Group 4 (10 atts. out of 10)		
'H'	44	88
'E'	40	65
'Other'	69	75

Table 4 shows the number of valid and invalid hypotheses advanced. Valid hypotheses were those tenable at the time they were advanced in the light of the information the subject had received. Valid hypotheses were very much more frequently advanced than invalid in all groups. Reducing the load on memory led to a significant increase in the number of valid hypotheses (sign test, $P < 0{\cdot}001$, two-tailed) but its effect on invalid hypotheses was more complex. In Group 1 and Group 2 the number of invalid hypotheses stayed at about the same level (sign test; n.s.) but in Groups 3 and 4 the number was reduced [sign test, $P < 0{\cdot}01$ (Group 3), $P < 0{\cdot}05$ (Group 4)]. (One exceptional subject in Group 4 advanced more invalid hypotheses than the rest of the group put together so means have been tabulated with and without that subject's results.) It is not surprising that Group 3 should have proposed fewer invalid hypotheses since in any case subjects in this group sought information about only those attributes that appeared to them to be relevant. The performance of Group 4 is however more remarkable as they had to cope with the whole mass of information and try to avoid being swamped by it.

Table 5 shows, for the first 11 trials, how many hypotheses were advanced involving 'H', 'E' and 'other' attributes. H and E were equally valid hypotheses but E was favoured in Groups 1 and 2 by being included in the initial attribute set. In these two groups E was proposed significantly more frequently than H ($\chi^2 = 26{\cdot}35$, $P < 0{\cdot}001$). In Groups 3 and 4, H and E were proposed with about equal frequency. 'Other' hypotheses were proposed more frequently in Groups 1 and 2 than in Groups 3 and 4.

DISCUSSION

The experimental task employed instances exemplifying ten attributes. Ten is not of course an indefinitely large number but it was considered large enough to be beyond the attention span of the subject and small enough to be manageable. Expt. I allowed information about four attributes of each instance (as in many orthodox experimental studies) but the form of the task made it apparent to the subject that there was a great deal of relevant information that was not immediately available to him. The instructions were that the predictor attribute might or might not be one of those about which information was given initially. Nevertheless subjects appear to have behaved as if it was; since when the predictor attribute was included in the initial set, 'guessed' solutions were significantly more frequent and a significantly larger number of attributes was left uneliminated. It is of course true that in real life the subject will include in his set of attributes to be observed any attributes he thinks *likely* to be the predictor. If he is lucky and one of his selected attributes *is* the predictor he will solve the problem quickly but, holding this view, he will be liable to arrive at a solution on the basis of inadequate evidence and the solution he arrives at may be wrong. An initial attribute set that does not include the predictor here predisposes the subject to caution. The same is true of knowing in advance whether the next instance will be positive or negative. Comparison between instances is the only way to eliminate irrelevant attributes and subjects who knew in advance were more likely to ask for further information about the same attributes thus maximizing their chances of eliminating one or more of them. It appears that the orthodox experimental task may be crucially unrepresentative because it capitalizes on the subject's inherent tendency to assume that his selected attribute set includes the predictor and on his tendency to gamble when he does not know in advance whether the next instance will be positive or negative.

Expt. II presented only 12 instances and since this was the minimum number in which a valid (non-guessed) solution could be achieved the subject's performance had to be evaluated rather than scored in terms of number of instances to solution. The comparison between minimum and maximum memory load conditions shows that subjects do very much better under minimum memory load. The number of 'logically correct' solutions increased significantly and the number of 'incorrect' solutions decreased. Each subject received exactly the same information in the two conditions (the information he had selected in the maximum load condition). Table 4 shows that a greater number of valid hypotheses was advanced under minimum memory load conditions. Performance under maximum load was, however, remarkably good considering that the subject never had more than one instance in front of him. A clue to the subject's capacity to work under these conditions (which provide the closest analogue to real life) may be found by comparing the performance of Groups 3 and 4. Group 3 subjects were allowed to select as many attributes as they liked of each instance and in general they selected all or nearly all of them for the first two or three instances and then only those attributes that appeared likely to be the predictor; three, two or even one per instance. Group 4 subjects were given information about all the attributes of every instance. It might have been expected that they would be swamped but their performance was like that of Group 3 subjects on every measure employed. It appears that Group 4 subjects did tacitly what Group 3 subjects were allowed to do explicitly, i.e. filtered out all information except that relevant to their current hypothesis or hypotheses. Group 1's performance was substantially worse than the rest and this may be attributable to the fact that subjects in this group received less information about the early instances in the series than they could usefully have employed.

The strategy adopted by Group 3 subjects (and believed to have been employed in all groups where the amount of information presented exceeded the subject's attention span) is of course what Bruner *et al.* called 'scanning', but it is seen here to be the strategy of choice wherever (as in real life) more information is available than can be put to use. 'Focusing' by contrast uses all the information derivable from a relatively small attribute set but can only be employed if the predictor attribute is known to be included in the set—this latter condition is fulfilled in the orthodox laboratory task. This result lends further support to the view that the orthodox task may be crucially unrepresentative.

The object of this paper was to see how far orthodox concept attainment tasks are representative of the corresponding real-world activity and the conclusion was that they are unrepresentative in crucial respects. In a recently published book (Levine, 1975) the author adopts a typically orthodox posture and a comparison between his practice and ours may be illuminating.

Levine writes (p. 289) 'This book contains a cheerful optimistic tale; after two decades of subterranean existence H-theory re-emerges victoriously, solving old problems with new insights and new techniques. It now dominates the areas of discrimination and concept learning and shows signs of incorporating neighbouring topics. Those of us associated with the theory can feel satisfied at having accomplished a durable piece of work.' He makes no reference to events in the real world outside the laboratory. What he sets out to explain are the results of his own and other peoples' laboratory experiments. Levine's scientific antecedents are impeccable. He began by extracting evidence for hypothesis (H) theory from primate data using elaborate mathematical techniques. He then developed and validated the 'blank trials' procedure for use with adult human subjects. (In this procedure instances are presented exemplifying four attributes each of which takes one of two values. A series of four instances presented without feedback enables the experimenter to determine objectively which of eight possible simple H's the subject is employing, or that the subject is not consistently employing any simple H.) He arrived at a characterization of human conceptual behaviour in the following terms: the subject selects his working H from a 'universe' of H's applicable to the type of problem he is required to solve, the universe may be partitioned into 'domains', i.e. one domain may consist of H's of the general form 'squares are positive, circles negative', another of H's like 'positive instances are in a sequence LLRLLR . . .' and so on. Given a positive instance the subject maintains his working H and at the same time evaluates other H's from the same domain (any H not exemplified by a positive instance must be wrong). Given a negative instance the subject drops his working H and selects another from his recollection of previous positive instances and at the same time (though not so efficiently as with a positive instance) evaluates other H's from the same domain (any H exemplified by a negative instance must be wrong). The subject learns nothing at all if he is selecting H's from the wrong domain and switches domains only if he exhausts a domain without finding a solution or in response to feedback about the correct domain.

This characterization describes fairly precisely what most of us thought we were doing anyway. The early behaviourists were of course well aware that this was the case but in their view 'what I think I am doing' is merely epiphenomenal and necessarily lacks causal efficacy. They adopted experimental paradigms that enabled them to study what they regarded as important – paradigms that represented the corresponding real-world activities in what they regarded as the theoretically crucial respects. One may legitimately use their experimental paradigms (as Levine does) but in logic one should then try to accommodate new experimental results by methods like those employed by, e.g. Spence (1936), to accommodate Krechevsky's results. If one comes to regard 'what I think I am doing' as more than an epiphenomenon, as possessing causal efficacy (as Levine also does), one has adopted a new metatheoretical stance and should do again what the behaviourists did in the first place; that is, clear one's mind of theoretical lumber and return (in Husserl's words) 'Zu den Sachen selbst' (to the things themselves), in order to construct a new experimental paradigm which is representative of the real-world activity in the respects *now* regarded as theoretically crucial. This is what the present authors have tried to do.

Levine gets the worst of both worlds. He employs an experimental paradigm in which pairs of very simple 'stripped-down' instances are presented and the subject knows that one will be positive and one negative. (This situation never occurs in the real world but was found helpful for rats in the early days of animal discrimination learning.) Even after he has shown (Karpf & Levine, 1971) that asking the subject what H he is using is as effective as the blank trials procedure, Levine continues to use the latter (he also adopts Skinnerian terminology and the subject's verbal response becomes an 'introtact'). And because intelligent adult subjects invariably use hypotheses in his experimental situation he feels bound to deny that they ever employ the more primitive procedures characteristic of rats, monkeys and children under five. Hull's (1920) results show that if attributes cannot be named and/or instances are presented too fast for proper processing the intelligent adult reverts at once to the more primitive procedure [see Wetherick (1970) for further discussion of this issue]. Incidentally, adult behaviour then conforms quite nicely to behaviourist theory.

More important are the limitations imposed on the possible future development of his theory by Levine's insistence on a basically behaviourist experimental paradigm. In his Epilogue he characterizes as 'a particularly pessimistic presentation' Dominowski's (1974) argument that different subjects employ different procedures in different situations – at the entire discretion of the subject, but he is unable to contest its essential truth. He goes on to propose two problem areas for future investigation. One concerns verbal coding. In Levine's situation the subject normally points to whichever of the two stimuli he thinks is positive and both are withdrawn before he is told whether he was right or wrong. His mental operations are then performed on his recollection of the stimulus he pointed to, whether or not it turned out to be positive. If, however, feedback is given three seconds before the stimuli are withdrawn the subject waits and performs his mental operations, not on the stimulus he thought was positive but on the stimulus that turned out to *be* positive. From Levine's point of view this behaviour is a matter for experimental investigation. From our point of view it is a matter of pure common sense on the part of the subject. The other problem area concerns more complex problem-solving situations but Levine's learning orientation here leads him to make no distinction between problems where the answer is a rule relating events in the real (or artificial, laboratory) world and problems where the answer is a procedure for bringing about a desired state of affairs. From our point of view this is an essential distinction. Levine's work may be germane to the former type of problem but not to the latter.

Thanks are due to Rosalind M. Heaton who ran the experiments and contributed to the development of the experimental procedure.

REFERENCES

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

DOMINOWSKI, R. L. (1974). How do people discover concepts? In R. L. Solso (ed.), *Theories in Cognitive Psychology: the Loyola Symposium*. Potomac, Maryland: Lawrence Erlbaum.

HULL, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychol. Monogr.* **28**, no. 1.

KARPF, D. & LEVINE, M. (1971). Blank-trial probes and introtacts in human discrimination learning. *J. exp. Psychol.* **90**, 51–55.

LEVINE, M. (1975). *A Cognitive Theory of Learning*. Hillsdale, New Jersey: Lawrence Erlbaum.

SPENCE, K. W. (1936). The nature of discrimination learning in animals. *Psychol. Rev.* **43**, 427–449.

WETHERICK, N. E. (1966). The inferential basis of concept attainment. *Brit. J. Psychol.* **57**, 61–69.

WETHERICK, N. E. (1970). Concept learning and the short-term memory – a contribution to theory. *Lang. and Speech* **13**, 271–278.

(*Manuscript received* 28 *July* 1975)

Br. J. Psychol. (1976), 67, 2, *pp.* 243–273
Printed in Great Britain

PARENT–OFFSPRING RESEMBLANCES IN INTELLIGENCE: THEORIES AND EVIDENCE

BY MICHAEL McASKIE AND ANN M. CLARKE
Department of Educational Studies, University of Hull

A review of data on parent–offspring resemblances in intelligence is presented in the context of correlation, regression and variance predictions from the polygenic model and an environmental model. Consideration was given to reliability, test equivalence and long-term stability of IQ scores. Results indicate that much of the difference among offspring IQ scores is not directly attributable to parental IQ, and that some is due to other between-family variables. The magnitude of single-parent–offspring correlations was related to the degree of assortative mating present in samples. Evidence on whether parent–offspring resemblances in IQ are transmitted genetically or environmentally is on the whole lacking. A single study provides evidence for some degree of genetic transmission, but it is argued that the present polygenic model is inadequate.

The problem of the contribution of heredity and environment to differences in intelligence is obviously not reducible to a single question. It is made up of a number of issues (e.g. wide and narrow heritability),* each of which may be investigated in a variety of ways. Thus a proper understanding of such a broad topic cannot be achieved in a single study of one aspect; an integration of evidence from numerous sources is required. However, the value of such an integration will depend crucially upon the soundness of each component contributing to it.

Several wide surveys have been published (e.g. Erlenmeyer-Kimling & Jarvik, 1963; Jensen, 1969; Jencks, 1973; Herrnstein, 1973; Eysenck, 1973), but if any of these reviewer's assessments prove correct (which they might), it will certainly not be due to the care with which they have handled the evidence. Even Jencks' most sophisticated attempt failed to match an elaborate mathematical analysis with a commensurate attention to the nature of the data.

This article is concerned with a detailed consideration of a small portion of the evidence: studies of parent–offspring resemblances in intelligence. Even within this narrow area, there are a number of ways of looking at the data. Firstly, the notion of resemblances can be interpreted as meaning either resemblance in actual performance or resemblance in relative level of performance with respect to an age group. For intelligence, the former requires a direct comparison between the performance of parents and offspring at the same age on the same test. Most studies use the latter method, the comparison being based on measures derived from performances on different sets of items at different ages. Unfortunately, psychological theories underlying test construction are inadequate to treat resemblance in type as meaningful. For example, the vagaries of item construction have led to the inclusion of different types of problem for young children and adults, despite the use of similar item analysis procedures. In these circumstances comparison is valid only to the extent that performance at one point in time relates significantly to performance at another. Thus the questions of correlations among tests and the predictive validity of IQ scores must be considered.

---

* Narrow heritability refers to the effect of the additive genetic components of the parents, whereas wide heritability refers to the effects of all genetic components, additive and non-additive. Parent–offspring correlational studies detect the additive component, the normal analysis of MZ/DZ twin studies picks up the wide component *within* (but not *between*) families.

Secondly, resemblances can be analysed by different statistical methods. Direct predictions can be made by regression equations and the extent to which such predictions are in error may be assessed by correlations. A correlation coefficient, of course, gives only an average error value and reveals nothing about the way in which errors are distributed. To determine this another kind of analysis may be needed. Whichever method is used, theoretical predictions can be made both for single-parent–offspring resemblance and for mid-parent–offspring resemblance. In addition, total population forecasts can be made.

In the following discussion of theoretical predictions, the assumption is made initially that differences in intelligence are due entirely to genetic factors or entirely to the environment. The fact that such positions are rarely, if ever, adopted does not vitiate the utility of considering their potential outcome.

GENETIC PREDICTIONS

The genetic predictions which have been made concerning parent–offspring resemblances have been entirely based on a simple additive polygenic model. On the whole the correlational predictions from this model are well understood; those for regression have occasioned some confusion, while those for variances have scarcely been contemplated. The failure to consider the latter predictions which are both easy to make and to test, represents a major fault in assessing adequately the predictions of the polygenic model. In view of psychologists' keen attention to sampling problems in other contexts, this appears to be a strange inconsistency.

Correlation predictions

The correlational predictions are well known, mathematically correct, and require no explication. With random mating the single-parent–offspring correlation is 0·5 and the mid-parent–offspring correlation is 0·7071. Despite this disparity, the two have been confused: thus, Erlenmeyer-Kimling & Jarvik (1963) include both single-parent– and mid-parent–offspring correlations in the same line of their well-known graph.

Assortative mating will increase both of these, with the single-parent correlation becoming closer to the mid-parent correlation as the degree of assortative mating increases. A problem in considering the effects of assortative mating on correlations (and on variance) is that they are cumulative over a number of generations (see Stanton, 1946). This is particularly pertinent to the genetics of intelligence where assortative mating is probably higher than for other characteristics.

Dominance will reduce the correlations, but since there is neither adequate empirical evidence, nor any theoretical reason, for assuming a particular degree of dominance, no meaningful predictions can be made. This lack of specificity has, however, been put to good use in accounting for anomalies in obtained results.

Regression predictions

The predictions concerning regression are somewhat less well understood than those concerning correlations. First, the relationship between correlation and regression does not seem to be fully appreciated, and secondly, single-parent and mid-parent relationships have again been confused (see Eckland, 1967; Eysenck, 1973; Herrnstein, 1973; Jensen, 1973, for examples). The confusion may have been helped by the lack of separate linguistic forms for referring to a parent taken singly and parents taken in pairs: the phrase 'relationship between parents and children' is totally ambiguous.

With random mating, single-parent–offspring regression will be 0·5 (i.e. regression to the mean), decreasing to zero as assortative mating increases to 100 per cent. *There is no regression from mid-parent IQ on the polygenic model.* The fact that the mid-parent–offspring correlation is less than unity (e.g. 0·7071) does not entail regression to the mean. Regression, in the usual sense of the word, is normally equal to 1 minus the regression coefficient from the regression equation. To deduce a regression coefficient from a correlation coefficient one must take into account the relative variance of the two variables [$y = r(\sigma y/\sigma x)x$] and it happens that the variance ratio effect exactly cancels the correlational effect for mid-parent–child regression (e.g. when $r = 0{\cdot}7071$, $\sigma x = 0{\cdot}7071\,\sigma y$, since mid-parent variance = half parent variance, and parent variance = offspring variance).

Variance predictions

Predictions can be made both about the pattern of variance of offspring IQ across different levels of parental IQ and also the total variance of the offspring compared with the total variance of the parents.

For mid-parent–offspring IQ the pattern prediction is dramatic. The variability of the progeny will be maximal when the mid-parent IQ is average and will be minimal (= 0) when the mid-parent IQ is extreme (high or low). Pairs of extreme parents will, of course, both be entirely isozygous* for high or low genes, a condition which can only lead to zero variance in the progeny. With random mating the single-parent–child variance will also decrease towards the extremes though to only half the maximum value (instead of zero as in the mid-parent case). With progressively greater degrees of assortative mating the single-parent pattern will increasingly resemble the mid-parent pattern.

The total variance of parents and offspring will be the same under random mating, but with assortative mating this will increase from generation to generation. The shape of the distribution will also change from normal to bimodal under assortative mating.

ENVIRONMENTAL PREDICTIONS

Before environmental predictions can be made an environmental model must be found, and this is not easy since most psychologists have been too canny to be explicit on the matter. However, implicit in studies of environmental effects there appears to be a simple additive vector model. It thus seems reasonable to see what the predictions of such a model might be for parent–offspring resemblances. This model of how the environment might operate is not one to which the authors subscribe. It is specific to two-parent families, and may not generalize to either single-parent or multiple-rearing conditions. The formulation was forced upon us by the total absence of any model, coupled with numerous vague and often potentially contradictory statements in the literature concerning parental environmental effects. It may, however, have value in indicating the direction in which reformulation may be necessary.

---

* Isozygous, meaning the genes have the same value, is used in preference to homozygous which additionally assumes identity.

Correlational predictions

If one were to assume that the environment produced by parental intelligence is the only relevant variable, that the effective contribution of each parent is equal to half his or her intelligence, and the contribution of both is the simple sum of the two, then, with random mating, the single-parent–offspring correlation would be 0·7071, and the mid-parent–offspring correlation, unity. Assortative mating would, of course, increase the former. These rather high correlations can easily be reduced to fit results by denying that parental intelligence is the only environmental variable, and attributing decreases to other factors incompletely correlated with parent intelligence.

Perhaps more realistically, the model can be adapted to make differential predictions concerning parent and/or offspring sex. Thus an 'exposure' model would predict mother–offspring correlations higher than father–offspring, assuming that mothers spend more time with children; an 'identification' model would predict mother–daughter or father–son correlations higher than mother–son or father–daughter. Neither of these patterns can be predicted on a sex-linked genetic hypothesis.

Differential predictions could also be made in relation to the period in which the environment is presumed to operate, *provided changes of environment are involved.* Since no study reported in this review concerns such changes, there is little point in considering the predictions.

However, it may be pertinent to observe that if an environmental variable operates cumulatively over time, correlations with it will not necessarily increase with age. Thus Jensen's (1974*a*) rejection of a cumulative environmental effect on the basis of failure of correlations to increase is invalid, since the particular context considered involved no change.

The same point applies to Bloom's (1964) conclusion that the first few years of life are environmentally critical since the longitudinal correlations change little thereafter. Furthermore, by contrast, genetic models can predict correlational changes with age, for while the genetic material is implanted at conception, it can become effective at any age, as in the case of hereditary diseases of late onset.

Regression predictions

Assuming parent intelligence as the only environmental variable, and that mating is random, single-parent–offspring regression half way to the mean would be predicted (the same amount, as with the genetic model). On an environmental model regression will decrease as assortative mating increases, and increase to the extent that the other variables, which determine the offspring's intelligence, are not entirely correlated with parent IQ.

The question of mid-parent–offspring regression is of particular interest since it has been fallaciously assumed that such an effect could only be accounted for by genetic influences. It has already been stated that the polygenic model predicts no regression from mid-parent; this fact is, however, frequently overlooked. Eysenck (1973) provides one example: in his collation of confusions about regression there is a genetic prediction using a formula in which the regression produced is actually determined by *what is not inherited*: Child IQ = Population mean + Heritability × (Mid-parent IQ − Population mean). In this case regression = 1 − Heritability. 'Heritability' is not equivalent to a genetic explanation, but merely states the extent of a genetic contribution. A total genetic explanation requires Heritability = 1; anything less than unity entails an environmental (including error) component. It is this *latter* factor which produces regression in the formula.

On the implausible assumption that parent intelligence is the only relevant environmental variable, the prediction concerning mid-parent–offspring regression is the same as that for the 'polygenic' model: none. Mid-parent–child regression, if it occurs, thus constitutes evidence against both models.

Variance predictions

The major issue concerning the variance pattern, in view of the somewhat surprising genetic predictions, is the distribution for different levels of mid-parent IQ. Any differences concerning single-parent–offspring patterns will be diluted versions of the mid-parent contrasts.

If parent IQ were the only relevant variable, then the offspring variance for different levels of mid-parent would be uniformly zero. If not, then the pattern of variance will depend on the distribution of other environmental variables with respect to mid-parental level. In the absence of any specifically expected non-regularities in the relative distribution of these, it may perhaps be supposed that they are evenly distributed. This 'in ignorance' model would then predict uniformity of distribution of offspring variance at each mid-parental level, whereas the polygenic model does not.

The implications of an additive environmental model for total population variance are theoretically interesting. Starting with the unlikely assumptions of random mating, exclusive influence of parent intelligence and equal parental contribution (amounting in deviation terms to half each parent's deviation from the mean) we find that the offspring population variance will be halved each generation. (Child variance will equal mid-parent variance, since child IQ = mid-parent IQ, and, with random mating, mid-parent variance will be half the single-parent variance.) Assortative mating will reduce the effect, but only if this were perfect, could the variance remain constant across generations. Variance reduction will also be less if environmental variables other than parent intelligence are operative.

There are only two ways to avoid an inter-generational variance reduction: either by postulating egression or by adding a variance-producing moderator variable. The first alternative would mean that the offspring IQ would exceed mid-parent IQ for bright parents and fail to reach it in the case of dull parents. Since empirically this is not the case, the second alternative becomes mandatory.

By a moderator variable we mean one which alters the effect of another without itself having any direct influence, and which also varies independently of the variable it modifies. The genetic model has, of course, such a 'variable' in the random segregation of parental genes. On this model only half of the total child variance is directly attributable to parental phenotype (predicted mid-parent–child correlation = 0·7071), the other half being due to the chance or capricious combination of genes. If a moderator were identifiable it would be correlated with outcome in the same way as an ordinary variable. The difference is that a moderator's effect depends in some way upon the values of the moderated variable. With a simple multiplicative relationship between the two, the contribution of the moderator variable would fall to zero when the variance of the moderated variable were also zero. By contrast the proportional contribution to variance of an independent variable would increase as the moderated variable variance decreased.

If moderator variables have to be introduced to maintain variance, their presence will also affect correlations, since the moderator variance contribution will constitute an error in the context of assessing the effect of the moderated variable.

Combining genetic and environmental models

Most psychologists are agreed that intellectual differences are not solely the product of either heredity or of environment, but of both. There would also appear to be substantial agreement that these two factors do not combine in a simple additive way: some kind of interaction is involved, though no one has as yet expressed precisely the nature of this interaction. It is for want of any properly formulated alternative that we have little choice but to consider the possibility of predictions based on an additive model in which only the proportions may vary. It is, of course, entirely possible that a collective judgement based on all these unspecified ways of interacting might in the end be additive.

The latitude present within each of the very simple models already outlined was sufficient to make it impossible to arrive at very much in the way of firm predictions. Plausible 'adjustments' of one kind or another could easily be made within either model. With these various latitudes now occurring in combination, perhaps the most precise prediction to be made is that the possibilities for 'adjustment' now become even greater.

Even on the assumption of an additive combination of genetic and environmental factors, few predictions are possible concerning the effect of different proportions of each; an obvious exception is the case of foster studies. When the problems arising from errors of measurement (as yet to be discussed) are added to these uncertainties, resolutions are even more difficult. In the circumstances one can do little other than suggest whether results fit better or worse one or other model.

For correlations, the restrictions resulting from the introduction of moderator variables make the environmental predictions very similar to their genetic counterparts in most cases. Clear distinctions only occur for the sex-related correlations. Regression is also non-discriminatory which leaves the neglected area of variance in which the clearest differential predictions can be made.

Table 1. *Test reliabilities*

Test	Reliability coefficients	Average
Alpha	0·94, 0·95, 0·90, 0·87, 0·91, 0·92, 0·98	0·93
Binet	0·83, 0·91, 0·91–0·97, 0·95–0·98	0·93
Cattell	0·80, 0·70s, 0·95	0·82
Otis	0·81, 0·96, 0·95, 0·87, 0·88, 0·93, 0·79	0·88
Raven's Matrices	0·79, 0·88, 0·79, 0·85, 0·83–0·93, 0·85	0·85
Wechsler/WAIS	0·97, 0·94	0·96

TEST ERRORS AND THEIR EFFECTS

The various theoretical predictions are based on the assumption that perfect measurement is involved. Now even if one accepts the definition of intelligence as that which is measured by intelligence tests and thus avoids all the problems of relating these measures to external criteria, this still leaves a number of 'internal' problems in interpreting test scores, namely: reliability, standardization, test equivalence and longitudinal stability.

Confronted with studies straddling an age span of 1–60+, conducted from the 1920s to the 1970s and employing a variety of intelligence tests, standardized in different places, and by no means self-evidently comparable, the possibility of test error as a major component in the variety of reported findings was apparent. It had been hoped that a careful analysis of these problems as they might pertain to studies of family resemblances would be available, so that a simple summary statement could be made here. Unfortunately our search proved fruitless, and we were thus forced to collect together data from a variety of sources in order to make at least some assessment of the order of magnitude of these potential errors.

Test unreliability

Table 1, which should be regarded as no more than a slight improvement on guessing, simply gives the reliability coefficients we have seen reported on the various tests involved in this article. It includes a mixture of split-half, test–retest and equivalent-forms reliability coefficients obtained on an assortment of samples and even involving different versions of the same test.

It seems not unreasonable to assume that test unreliability will operate largely as a random error, entailing a reduction in parent–offspring correlations of something between 4 and 18 per cent depending upon the test involved. It is possible that the quoted reliability figures are slight underestimates, involving as they do additional 'errors' of delay and/or equivalence. For this reason correction for attenuation due to test unreliability is regarded as a dubious procedure. Since the errors here are probably small, and there is little evidence of consistent differences with respect to reliability from test to test, it seems unlikely that this factor will vary much among studies.

Test-standardization errors

Test-standardization errors in the form of gross differences in means and s.d.s *between* parents and offspring test scores will, of course, produce errors in regression and population variance predictions but correlation would be unaffected. Errors *within* each group (parents or offspring) will produce an unsystematic effect similar to test unreliability. In almost all studies both parents and offspring are each composed of groups of different ages and thus liable to such within-group effects if the means and s.d.s at each age are not the same.

For children aged 5–18 years, the 1937 Stanford–Binet standardization data (McNemar, 1942) show that the s.d.s vary 1·3 points, on average, about the mean of s.d.s and the means vary 1·2 points. Crude calculations using these average errors reveal that they would only reduce correlations by ½ per cent. It is unclear whether the above are normative errors or sampling errors. Dearborn & Rothney (1941) have shown that sampling differences occur between tests, but the extent to which they occur within tests can only be a matter of speculation; however, even mean and s.d. disparities of 10 points produce only an 8 per cent error.

For adults, cross-sectional samples (which are the appropriate ones in this context) reveal appreciable changes in mean IQ according to age. The typical adult sample in parent–offspring studies has an average age of between 40 and 45 years and a range of about 40 years. A crude estimate based on the data of Wechsler (1958), Doppelt & Wallace (1955) and Schaie & Strother (1968) suggests that the average disparity within such a range will be about 10 points. For s.d. changes with age we estimate an average disparity of about 3 points, based solely on the *graphs* of Foulds & Raven (1948). Combined, these errors would produce a 5–6 per cent reduction in correlations when age corrections have not been applied.

Alström (1961) and Cattell & Willson (1938) provide comparisons of age corrected and uncorrected correlations; for the former the differences are 6, 13 and 4 per cent, and for the latter 23 per cent. The effect of using two tests with disparate means on the children in Outhit's (1933) study was looked at, but here the rather small predicted errors (2·8, 2·4, 1·6 per cent) did not fully materialize (2, 0·7, 0·15 per cent).

Test equivalence

No less than six different intelligence scales have been used in the studies included in this review; furthermore, various forms of a single scale may occur not only across studies, but also within them, e.g. the 1916 and 1937 Binet (Forms L and M). Obviously an important issue is to what extent these various tests are measuring the same thing.

The prediction of changes in parent–offspring correlation resulting from errors due to lack of test equivalence depends upon the error model adopted. Assumptions have to be made on two issues: firstly, the location of the errors, and secondly, whether they will operate randomly or systematically across generations. On the first issue (location), any lack of correlation between two scales could be interpreted either as error in only one of them (i.e. Scale 1 measures Trait A and Scale 2 measures Trait A plus error) or as error in both (i.e. Scale 1 measures Trait A plus error; Scale 2 measures Trait A plus error). Where there is error in only one scale, a change of scale must involve a change of error variance and thus a change in correlation is expected. However, where there is error in both scales, a change of scale simply involves substituting one error for another, which, given error equality, produces no change in correlation.

On the second issue, a problem arises when different tests appear in two or more

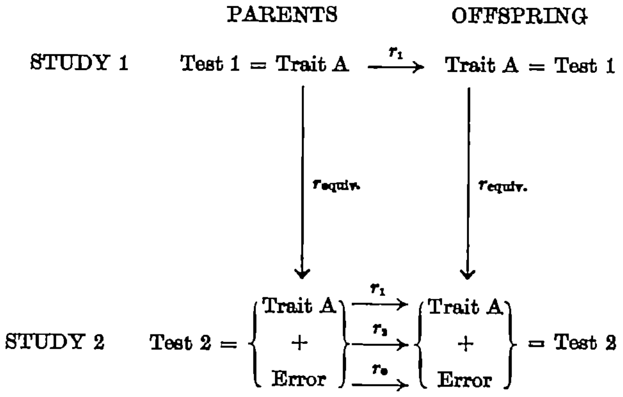


Fig. 1. Diagram showing the operation of a test equivalence error across two studies.

Table 2. *Inter-test correlations*

	Binet	Cattell	Otis	Raven's P.M.	Wechsler (WAIS & WISC)
Alpha	0·73, 0·8–0·9 Av. = 0·81	—	—	—	0·74
Binet		0·56–0·85 Av. = 0·71	0·63, 0·51 Av. = 0·57	0·71, 0·86, 0·62 0·54, 0·68 Av. = 0·68	0·83, 0·91, 0·62, 0·93 0·89, 0·82, 0·76, 0·77 Av. = 0·82
Cattell			0·73	—	0·84
Otis				—	0·67
Raven's P.M.					0·91, 0·62, 0·85, 0·75 0·55, 0·72 Av. = 0·73

studies. In the simplest case shown in Fig. 1, Study 1 uses Test 1 for parents and offspring, while Study 2 uses Test 2 for both.

Here, the error arising from the lack of equivalence of Test 1 and Test 2 might either be random across parents and offspring, or constitute a trait showing a definite relationship (r_e). In the former case it functions as an intergenerational error further reducing the correlation, and in the latter it constitutes a confounded variable which, if $r_e > r_1$, would increase the correlation.

In view of these indeterminancies, precise predictions cannot be made and one can only specify the upper limits of such errors from test equivalence data. Table 2 shows our collection of test intercorrelation data which once more, hopefully, represents an improvement on guessing.

The upper limit for errors resulting from imperfect test equivalence would be a reduction of $(1-r_{\text{equiv.}}) \times 100$ per cent in the parent–offspring correlations. With the overall average $r_{\text{equiv.}}$ from the Table (0·73) this would amount to a 27 per cent correlation reduction. Test equivalence, or rather lack of it, thus constitutes a potentially large source of error.

Relevant evidence on the actual magnitude of these errors is both scant and in varying degrees indirect but fairly encouraging. Skodak & Skeels' (1949) study shows an 8 per cent disparity between the 1916 and 1937 Binet correlations (estimating the reliability loss in this study is not possible since testing was done in a single session with items common to both tests being taken only once). Conrad & Jones (1940) and Outhit (1933) used the Army Alpha for adults and older children and the Binet for younger children, thus providing a same/different comparison confounded (adversely) with age. The disparities (for single-parent–offspring correlation) are 0 and 6 per cent respectively (7 per cent mid-parent). Indirect evidence from longitudinal studies of child IQ involving test switches between Binet and Wechsler variants, and in one case the Terman Group Test, produced no test change effect (Skodak & Skeels, 1949; Kangas & Bradway, 1971; Honzik *et al.* 1938; Pinneau, 1961). Unfortunately we have no evidence at all on this matter for the Cattell and Otis tests nor for Raven's Matrices. However, the Otis and Army Alpha appear fairly similar, which is not surprising in view of their historical connexion.

Longitudinal stability

Earlier we commented that in the absence of an adequate intelligence test theory scores obtained at different ages could not meaningfully be compared directly, but only to the extent that one was predictive of the other. This issue is thus effectively a special case of test equivalence and contains the same kind of error prediction indeterminancies, i.e. one can only estimate the maximum possible disparities. Direct evidence for estimating these is mostly lacking for the age differences between parents and offspring to be found in the studies. Hence, we attempted to make rather crude extrapolations from the longitudinal data provided by: Bayley (1949, 1957), Bradway & Thompson (1962), Ebert & Simmons (1943), Hirsch (1930), Honzik *et al.* (1938), Honzik (1972), Kangas & Bradway (1971), Sontag *et al.* (1958).

Our somewhat *ad hoc* procedure consisted of grouping the data into age bands (e.g. early, middle and late childhood) and combining the results in each age band using weighted z scores. The data were then plotted graphically (see Fig. 2), spacing the age bands at equal intervals along the age scale and carrying out linear extrapolation, weighting the projection according to the ns at each point in the one case where this was applicable. Clearly the validity of our extrapolation depends upon the appropriateness of our age scaling. The reasonable uniformity of the slopes of the different plots (each representing a different starting age) suggest that the procedure was not radically in error.

Since the average age of most parents in the studies falls between the middle and late adult bands, we have computed the longitudinal correlations to this point from each of the earlier bands (see Table 3). Given our estimates of longitudinal instability, it would be possible for parent–offspring correlations to increase or decrease by as much as 18–61 per cent, depending on the age disparity involved.

Test errors in combination

If parents and offspring were entirely homogenous with respect to age, and were given the same test, then only errors of reliability would operate. If one assumes that these are randomly distributed, they would reduce parent–offspring correlations by about 5–10 per cent. Regression would be increased proportionately, and overall variance trends masked to this extent.

Since almost without exception the studies include parents and offspring who vary

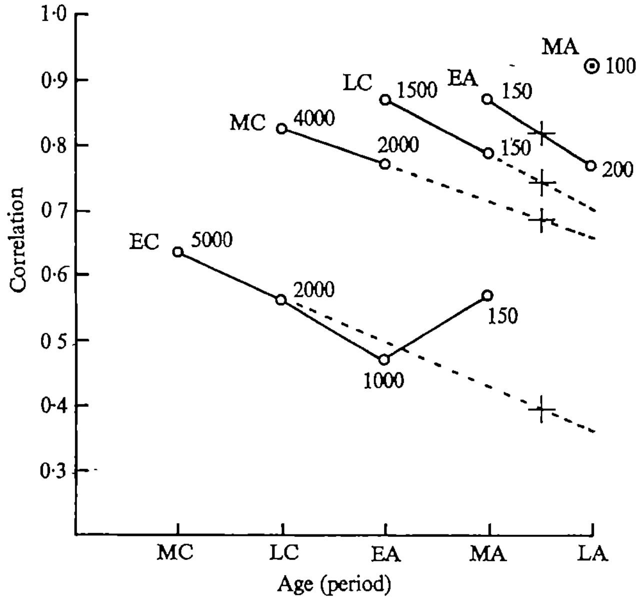


Fig. 2. Graphs showing longitudinal correlations for IQ scores from different age bands together with projections. Numbers beside points denote the number of entries in the correlation. EC, Early childhood (3–5 years); MC, middle childhood (6–9 years); LC, late childhood (10–14 years); EA, early adulthood (15–21 years); MA, middle adulthood (22–42 years); LA, late adulthood (43 years+). Points marked + are used for estimates in Table 3.

Table 3. *Estimated longitudinal correlations to middle/late adulthood* (40–45 *years*)

From early childhood (3–5 years)	0·39
From middle childhood (6–9 years)	0·69
From late childhood (10–14 years)	0·74
From early adulthood (15–21 years)	0·82

around a mean age, standardization errors will also operate, again presumably randomly, and to the extent of about 5–10 per cent. Combining the two, an expected reduction in correlation of between 10 and 20 per cent should be envisaged.

However when one considers the less than perfect equivalence of tests, and the facts concerning long-term changes of test score, the added uncertainty concerning the error model means that one can do little more than estimate the upper and lower limits of their effects on correlation.

Test equivalence errors were placed as potentially ranging from 0 to about 30 per cent. Since the assessment of test equivalence involves test unreliability already, there is no need to add this again to the maximum estimate, but it does need adding to the minimum, making the range from 5 to 30 per cent. Combining this with standardization errors gives a range between 10 and 40 per cent for studies involving a change of test.

Longitudinal errors were placed between 0 and 60 per cent but the latter figure was for prediction from early childhood. An upper region of 20–30 per cent is most applicable for most studies. Again longitudinal errors already include unreliability hence their effect, when combined with standardization errors, could be between 10 and 40 per cent.

Combining the effects of equivalence and standardization errors involves unreliability twice, for which allowance must be made. Furthermore, while it is reasonable simply to add small errors, this obviously should not be done for larger ones: each error must be applied to what the preceding errors have left. Our estimate, based on the speculations enumerated above, is that all test errors in combination may affect parent–offspring correlations minimally to the extent of 10 per cent and maximally 65 per cent.

THE STUDIES

An attempt has been made to produce a comprehensive collection of studies which include data on natural parent–offspring resemblances in intelligence. Those included all meet the following criteria: that intelligence tests were given to both parents and offspring, reasonably appropriate analyses carried out, and the population sampled not grossly abnormal. For each paper a brief outline is provided, giving some detail concerning the size and nature of the sample, the country or origin where outside the USA, the tests used, and such corrections carried out by individual authors as feature in the data we have used. Statistical data relating to these investigations are collated in Table 4. The studies are presented in rank order of size of single-parent–offspring correlation, both in the outlines and in Table 4.

Cattell & Willson (1938)

Sample. 101 English families, initially those of a test standardization sample willing to cooperate, resulting in a skew to high IQ. To compensate, parents of schoolchildren in classes for the dull and two guidance clinics were approached. Markedly bimodal distribution of parents. Range – corrected correlations used.

Tests. Cattell Scales 0 to III, designed to have an unusual s.d. of 27. Age corrections applied to older scores.

Outhit (1933)

Sample. 51 families of 4 or more offspring, most American born, all native English speakers, literate. Excluded: mentally ill, alcoholics, epileptics, TB, VD and unwillingness to cooperate. Biased to high SES.

Tests. Parents and offspring over 12, Army Alpha. Offspring below 12, 1916 Binet. Several growth limits for the Army Alpha were tried: the first yielded appreciably higher IQs of offspring than parents. Age 14 years and 10 months was chosen, giving same average of parents and offspring, to fit Outhit's assumption of no difference (pp. 21 and 22). Total data on age and IQ included; reanalysed by us; minor discrepancies; only our figures included.

Leahy (1935)

Sample. Control group for a foster study, 194 families, of North European, non-Jewish extraction, in urban communities of over 10000. Markedly skewed to high SES. Only one child per family tested.

Tests. Parents, Otis Intermediate A. Offspring, 1916 Binet. Supplementary data give a mid-parent–offspring correlation for natural children of some foster parents.

Conrad & Jones (1940). An extension to a study published by Jones in 1928.

Sample. 269 families, representative of rural New England population. All families native-born, English speakers, literate, with two or more offspring. Original sample shows negative skew in distribution of parents' test scores.

Tests. Parents and offspring over 10, Army Alpha. Offspring age $3\frac{1}{2}$ – 13, 1916 Binet. Between ages 10 and 14, 44 offspring were tested on both and a composite score used. Results converted to sigma scores, based on data from 1215 people aged 10–65 tested in the same community.

Alström (1961)

Sample. 50 families selected from a sample of twin pairs, resident in Greater Stockholm, Sweden, first identified at induction to military service, and willing to cooperate. Offspring: 26 monozygotic, 24 dizygotic pairs, plus 82 siblings, virtually all adult. Parents said to be a representative sample of Swedish population.

Test. Swedish Wechsler (excludes: Digit Symbol, Block Design and Object Assembly). Corrections made to scores over age 37. Means and s.d.s appear based on uncorrected scores, but correlations on corrected scores. Mid-parent–offspring correlation: personal communication by Professor Alström. In calculating parent–offspring correlations, mid-twin IQ was used, each pair one entry only.

Burks (1928)

Sample. Control group for a foster study; 105 families representing wide geographical area: all English speaking; markedly skewed to high SES. Only one child per family was tested.

Test. Slightly shortened 1916 Binet (both parents and offspring). Considered age correcting scores but decided against.

Higgins et al. (1962)

Sample. Part of a cohort numbering 82217 collected as follows: The grandparents of 289 residents in a colony for the mentally retarded sometime between 1911 and 1918 were traced. Information on descendants and their spouses, sometimes through seven generations, was obtained. Of these only 2·7 per cent were mentally retarded, so intellectual normality is assumed. This study reports on those families for whom IQ values were available for both parents and one or more of their offspring (1016 mothers, 1016 fathers and 2039 offspring).

Tests. Taken from various records, mostly school. Group tests of the Kuhlmann or Otis type, also some unspecified individual tests. Mean age of parents at test, 14·2; offspring, 8·65.

Kagan & Moss (1959)

Sample. Ninety-nine children from the Fels longitudinal study ($n = 140$), tested regularly from age 2½ to 11. Approximately 50 per cent of fathers and 27 per cent mothers were college graduates: high SES.

Tests. Mothers, Otis A; fathers, Otis B. Offspring, 1937 Binet, forms L and M alternatively. 'Smoothed IQs': average of three scores around each age reported.

Gibson (1970)

Sample. Scientists below the age of 35, employed by Cambridge University, England ($n = 157$, mean IQ 126·5). Of these 76 had fathers who were tested, 71 per cent were from the two highest social classes.

Test. Fathers and Scientist sons: WAIS. Since this study is based on *offspring* of known high academic achievement, the correlation is eliminated from most analyses.

Waller (1971). See under Higgins *et al.* for description of original population and tests.

This sample. One hundred and seventy-three male offspring aged 24 or older, and 131 fathers. Mean age of fathers at test: 15·9; sons 13·38. Degree of overlap, if any, between this sample and that used by Higgins *et al.* is unknown.

Polansky et al. (1969)

Sample. Young children in day-care centre with low IQ mothers.

Tests. Mothers, WAIS. Offspring, Binet.

McCall (1970)

Sample. Parent–offspring pairs selected from the Fels longitudinal study (see Kagan & Moss for sample and tests). Thirty-five pairs in which parents and offspring were tested at the same age in childhood.

Guttman (1974)

Sample. One hundred Israeli families in middle-class neighbourhoods. Of 200 parents, 131 were doctors, lawyers or teachers, mostly from Eastern Europe.

Test. Raven's Progressive Matrices (individual, 35 min). Age range of sample taking the same test was 8–61; all scores age corrected, but procedure seems dubious. Correlations calculated on scores: no IQ conversions.

Skodak & Skeels (1949)

Sample. Part of a sample of 100 children from below average social background placed in adoptive homes under the age of six months. In 63 cases the natural mother's IQ was known.

Tests. Mothers, 1916 Binet. Offspring, 1916 Binet at mean ages 2, 3, 7 and 13 (approx.). Also 1937 Binet at 13. Since all data are included in the original paper, we reanalysed parent–offspring correlations.

Oden (personal communication)

Sample. Part of the Terman gifted group (n = 1528) selected in childhood as having IQs above 135. As is well known, in later life, regression to the mean had occurred (McNemar, 1947). The correlation reported is on 559 gifted subjects tested as children, and 1027 offspring.

Tests. Parents, 1916 Binet. Offspring, 1937 Binet, Form L.

Snygg (1938)

Sample. Canadian cohort of 312 young children of low social background placed in foster homes (79 per cent in first year of life, 90 per cent before 2).

Tests. Mothers, 1916 Binet. Offspring, Kuhlmann–Binet and 1916 Binet.

Freeman et al. (1928)

Sample. Part of a large study of foster children. In 28 families a comparison was made of foster parents' IQ with their own and foster children. Only mid-parent–offspring correlation reported.

Tests. Parents, Otis Higher Form. Offspring, 1916 Binet.

Jones (1928). See Conrad & Jones.

A number of studies which did not meet the criteria for selection were excluded:

Adams & Neel (1967) – offspring were the result of incestuous relationships.

Brandon (1957) – various tests used, including social maturity; and fathers' IQs were estimated.

Burt (1943, 1955, 1961) – the 1955 study used estimates of parental IQ. The 1961 study is probably the most widely quoted statement on parent–offspring regression, and claims to develop the 1943 data. However, careful inspection reveals a grossly inadequate description of the material and how it was collected. References to further details in other papers proved empty, and the trail through a number of publications led nowhere. Despite some cautious phrases, including one on the relative unreliability of the adult IQ assessment, the 1943 paper proclaims a mean IQ of 153·2 for his higher professional group, which exceeds that of the Terman gifted group, at a time when Terman could not find a test suitable for such exceptional adults. By 1961 the mean IQ of the higher professional group had miraculously shrunk to 139·7, but, surprisingly, the range for the offspring underwent no change during the same period. Several authors have recently attested to the unsatisfactory nature of Burt's work (Clarke & Clarke, 1974; Jensen, 1974*b*; Kamin, 1973). The latter has concluded that 'the numbers left behind by Professor Burt are simply not worthy of serious scientific attention'.

Eaves (1973) – this analysis of the Reed & Reed (1965) data was excluded for using analysis of variance which is inappropriate to the question of parent–offspring resemblances.

Halperin (1945) – some IQs were estimated, and all the data categorized for computing the correlation.

Honzik (1957) – parents' IQs were estimated from education.

Penrose (1933) – some IQs were estimated and the correlation based on categorized data.

Willoughby (1928) – a curious psychometric cocktail of subtests from various sources was used, hardly to be regarded as a standardized intelligence test.

DATA ANALYSIS

In analysing the data, we have been concerned to look at all the variables which might separately, or in combination, be relevant to the question of parent–offspring resemblance, regardless of their potential significance to a genetic or an environmental model as at present conceived. In doing so we have borne in mind the possibility that neither model is adequate, and that the empirical results available have been hard obtained, thus deserving of study in the hope that new hypotheses might emerge. An attempt has been made to carry out a relatively systematic analysis of the data dealing in turn with the correlational, regression and variance findings, looking for both consistencies and anomalies in the results.

There were a number of statistical problems in handling the data. The two basic sampling problems require different treatments when combining the results of different studies. The within-study sampling problem, i.e. the extent to which the data from a study is representative of the sampling procedure used within it, depends upon the n in the study and is dealt with by using weighted z scores,

$$z_{\mathrm{av}} = [\Sigma\, z(n-3)]/[\Sigma\, (n-3)].$$

However, errors due to different sampling procedures between different studies are unaffected by n and best dealt with by taking unweighted z scores. With such a procedure the problem arises as to what n to use for significance testing. To test for non-significance we have treated every study as having the n of the largest study, and testing for significance have used the smallest n, thus in each case making a conservative estimate.

The appropriate n for each study also presents problems because the same individuals may appear more than once in the correlation calculation. Again we have used maximum or minimum values according to which was the least favourable for the hypothesis being tested; the maximum value being the number of entries. Where the proper minimum value could not be established, we have used the number of families since the minimum cannot be less than this.

For direct comparisons, all the various correlations should have been corrected or not to the same degree, but unfortunately almost every study seems to have gone its own way in this respect. Since the majority of correlations were not age corrected, we have taken non-age corrected values where available. None of the correlations include corrections for attenuation. We considered carrying out range corrections since the s.d.s varied from study to study, but were unable to establish the appropriate s.d.s for every test. For example it was impossible to adjust Burks' correlation derived from non-age corrected 'lopped' Binet IQs.

Readers should be warned that some of the figures quoted in various tables and used in our analyses have been derived from published data, but do not always correspond exactly with those their authors chose to present, or those cited by others. In some cases our reanalysis of the published raw data yielded marginally different figures; in others it was possible to calculate from the data given results which were

Table 4. *Summary data on studies of natural parent–offspring resemblance*

	Correlations			*n*s		Mean IQ		s.d.		Average age		Tests		
Author and date	Single parent–offspring	Mid-parent–offspring	Assort. mating	Families	Offspring	Parents	Offspring	Parents	Offspring	Parents	Offspring	Parents	Offspring	Additional information
attell & Willson (1938)	0·80	—	0·71	101	101	106	109	32*	29*	Adult	Child	Cattell	Cattell	* s.d. for test: 27
uthit (1933)	0·62	0·67	0·74	51	257	107*	107*	19	17	42	12	Army Alpha	1916 Binet Army Alpha	* Means made equal
eahy (1935)	0·51	0·60	0·41	194	194	—	110	—	15	Adult	9	Otis Intermed.	1916 Binet	—
onrad & Jones (1940)	0·49	—	0·52	269	506E	−0·06*	0·15*	1·01*	1·06*	41	8 and 16	Army Alpha	1916 Binet Army Alpha	* Sigma scores
lstrom (1961)	0·49	0·56	0·47	50	132	99	106	14	13	56	28	Wechsler (Swedish)	Wechsler (Swedish)	—
urks (1928)	0·45	0·52	0·55	105	105	111	115	19	15	39	8	1916 Binet (lopped)	1916 Binet (lopped)	—
iggins *et al.* (1962)	0·44	—	0·33	1016	2039	102	106E	15	16E	14	9	Various 'Otis type'	Various 'Otis type'	—
agan & Moss (1959)	0·43 Av.	—	—	> 59	99	Above av.	Above av.	—	—	Adult	3–10	Otis	1937 Binet (smoothed) possibly some 1916s	—
ibson (1970)	0·37	—	—	76	76	123	131	11	6E	Adult	Adult	WAIS	WAIS	—
Valler (1971)	0·36	—	—	131	173	99	103	16E	15E	16	13	Various 'Otis type'	Various 'Otis type'	—
olansky *et al.* (1969)	0·35	—	—	65	65	78M	94M	—	—	Adult	5	WAIS	Binet	—
IcCall (1970)	0·32 Av.	—	—	35	35	—	—	—	—	3–11	3–11	1916 and 1937 Binet (smoothed)	1937 Binet (smoothed)	—
uttman (1974)	0·28	0·41	0·26	100	208	110E	110+E	—	—	48M	16M	Raven's P.M.	Raven's P.M.	—
kodak & Skeels (1949)	0·27 Av.	—	—	63	63	86	113Av.	16	14Av.	Adult	2–14	1916 Binet (mostly)	1916 and 1937 Binet	—
den (person. comm.)	0·08	—	—	559	1027	152	133	15	18	Child	Child	1916 Binet (mostly)	1937 Binet	—
nygg (1938)	0·05 Av.	—	—	312	312	78	95	11E	10E	Adult	1–5+	1916 Binet	Kuhlmann–Binet 1916 Binet	—
reeman *et al.* (1928)	—	0·35	—	28	28	—	—	—	—	Adult	10E	Otis Higher	1916 Binet	—
eahy (1935) additional	—	0·36	—	20	20	—	—	—	—	Adult	—	Otis Intermed.	1916 Binet	—
ones (1928)	—	0·59	0·60	105	—	—	—	—	—	—	—	Army Alpha	1916 Binet Army Alpha	—

Notes: M, median; E, estimate; Av., average for same group at different ages.

not pertinent to the original authors' argument, but which were needed for our purpose.

RESULTS

Single-parent–offspring correlations

Casual inspection of the column of single-parent–offspring correlations in Table 4 reveals a range from 0·80 to 0·05 (subgroups from two studies, Snygg and Gibson, are reported with small *negative* correlations). Needless to relate, no theoretical model predicts such diversity, hence a major consideration in our analyses was attempting to account for it.

Sex of parent and offspring

Four studies report separate correlations for father–son, father–daughter, mother–son and mother–daughter (Conrad & Jones, Guttman, Kagan & Moss, Outhit). Despite large numbers of entries, comparisons yielded nothing approaching significance. Similarly, no difference was obtained when comparing father–offspring with mother–offspring correlations (Alström, Burks, Cattell & Willson, Conrad & Jones, Guttman, Kagan & Moss, Leahy, Outhit, Higgins *et al.*), this time with even larger *n*s. Thus, all such distinctions have been ignored, where separate correlations were given they were collapsed, and studies giving only single-sex parent–offspring correlations treated as equivalent.

Dominance by superior versus inferior parent

Two authors examined a possible differential relationship between offspring IQ and either the more or less intelligent of each pair of parents. The results were essentially negative: for Jones the correlation with the superior parent was 0·57 and for inferior 0·55; Outhit reports 0·57 and 0·62 respectively.

Age: parents tested as children versus parents tested as adults

Excluding foster studies and deviant IQ group studies (to be discussed), there were three investigations in which the parents were tested as children (Higgins *et al.*, McCall, Waller) and eight in which the parents were tested as adults (Alström, Burks, Cattell & Willson, Conrad & Jones, Guttman, Kagan & Moss, Leahy and Outhit). The former group produced lower correlations than the latter as a group: weighted $r = 0{\cdot}43$ and 0·52 respectively, $P < 0{\cdot}01$; unweighted $r = 0{\cdot}37$ and 0·52 respectively, $P < 0{\cdot}08$. Since only three studies involved the parents as children (two from the same sample and possibly overlapping) and since the unweighted significance is low, we are disinclined to accept this evidence as more than suggestive.

Age of offspring at testing and within-study test changes

Two studies are commonly cited as showing an increasing parent–offspring correlation as the age of the latter increases: Honzik (1957) and Skodak & Skeels (1949). The former used estimates of parental IQ from education; since this procedure probably involves more 'error' than 'effect', the evidence must be regarded as weak. In the latter study (of fostered children) there does appear to be a rising correlation in early childhood. However, inspection of Fig. 3 makes it apparent that neither other

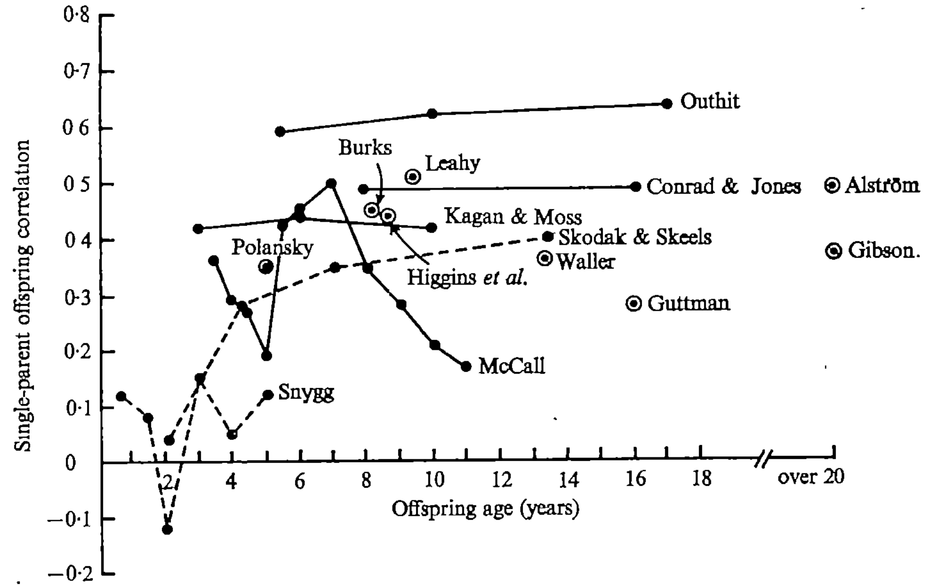


Fig. 3. Graphs showing single-parent–offspring correlations by offspring age. – – –, Foster study.

longitudinal studies (Snygg, McCall, Kagan & Moss) nor cross-sectional studies (Outhit, Conrad & Jones) show a similar trend. Nor is there any evidence for an increase in parent–offspring correlation beyond the age of four across the studies. However, it remains possible that there is an age trend from birth to about four years.

It should be noted that both Outhit and Conrad & Jones used a different test for younger children than for both older offspring and adults. Here we have, therefore, a change of test *added* to the age change, yet with no significant effect. Although in about half the studies involving young children, different tests were used for parents and offspring, above the age of 14 (offspring) the same test was used in all cases for both, again without apparent effect. It is difficult to avoid the conclusion that, on the whole, above the age of four, neither age nor the type of test used systematically affects the parent–offspring correlation, and to note that the studies by Outhit and Conrad & Jones using virtually the same procedures with respect to tests and age, yield substantially different results.

Fostering and non-fostering studies

If there is any environmental contribution to offspring intelligence, a diminished natural parent–child correlation would be expected in foster studies. There are only two such reports in the sample (Snygg, Skodak & Skeels). In addition to the problem of few studies, there are two further complications: firstly Snygg's involves fairly young children, and secondly the mothers in both groups are on average of low IQ.

Only the most stringent comparison between foster and non-foster studies (in the latter case excluding Oden and Gibson as deviant groups) failed to reach an acceptable level of significance on a one-tailed test: $P < 0{\cdot}07$. This particular comparison involved taking correlations with offspring above age four only, using unweighted

Table 5. *Single-parent–offspring correlations for deviant IQ groups*

Study	r	Parent dev. from 100	Parent IQ Mean	Parent IQ s.d.	Offspring IQ Mean	Offspring IQ s.d.	n of offspring
Oden	0·08	+52	152	15	133	18	1027
Gibson	0·37	+23	123	11	131	6	76
Snygg	0·13	−22	78	11E	95	10E	312
Polansky *et al.*	0·35	−22	78M	—	94M	—	65
Skodak & Skeels	0·38	−14	86	16	106	15	63

Notes: Gibson, means computed from table of cases used in correlation; Snygg, s.d.s estimated from s.e. of mean; Polansky *et al.* medians only given; Skodak & Skeels, IQ data as given at age 13 (1916 Binet). E, Estimate; M, median.

averages, with minimum n, and range-correcting Snygg's correlation from 0·13 to 0·28, a dubious procedure based on s.d. estimates which may well have been too low. All other comparisons yielded differences significant beyond the 0·02 level.

Thus, taken as a group, the correlations in these two studies appear to be lower than others, but this ignores the obvious disparity between the two (see Fig. 3). It is curious that Snygg, with an individually tested sample of 312 offspring and their mothers, is so rarely cited.

Deviant IQ groups

Five of the studies concern parent groups which deviate about one s.d. or more from 100. Table 5 gives the relevant details.

The average correlation of this group is significantly lower than the average of less deviant groups ($P < 0{\cdot}001$). Two of the five studies have already been discussed as involving fostered children. Other complications arise: (1) with the exception of Gibson and Skodak & Skeels, a number of young children are included; (2) Oden's correlation is based on parents tested as children; (3) the added unreliability of assessing high IQ; and (4) Gibson, certainly, and possibly Snygg have low s.d.s.

The Terman–Oden correlation has not previously been published, and was made available through the kindness of Dr Oden, who wrote: 'The number of Binet-tested parents in the chart is limited to those for whom offspring Binet IQs were also available in 1955...The correlation between parent and child test scores was not published in *The Gifted Group at Mid-Life* (vol. V of the series) because the restricted range made it impossible to get a high r. The obtained r is significant only because of the large ns.'

Assortative mating

Inspection of the columns in Table 4 giving single-parent–offspring and parent–parent correlations indicates that there is considerable variation in the latter, and that the two are related. Table 6 presents the available data in rank order of assortative mating. Even on a simple product-moment correlation, the relationship between single-parent–offspring r and assortative mating is significant: $r = 0{\cdot}85$, $P < 0{\cdot}01$.

An attempt was made to remove the effect of assortative mating from the single-parent–offspring correlations using the partial correlation procedure. This involves

Table 6. *Assortative mating and single-parent–offspring correlations*

Study	Assortative mating	Single-parent–offspring	Single-parent–offspring (partial)
Outhit	0·74	0·62	0·31
Cattell & Willson	0·71	0·80	0·55
Burks	0·55	0·45	0·27
Conrad & Jones	0·52	0·49	0·32
Alström	0·47	0·49	0·34
Higgins *et al.*	0·33	0·44	0·35
Guttman	0·26	0·28	0·22

considering the parents as separate but correlated variables. The results are included in Table 6. In view of the ensuing considerable reduction in the variability of correlations, it seems reasonable to conclude that a powerful source of difference between the separate studies is the degree of assortative mating. It will, however, also be noted that, with one exception, all the correlations are reduced well below that expected by either a genetic or an environmental model (as at present conceived), granted random mating, and assuming that parent IQ related variables are of major importance.

Mid-parent–offspring correlations

It will be recalled that the genetic model predicts a mid-parent–offspring correlation of 0·7071 and the simple parent IQ related environmental model a correlation of 1·0. Column 2 of Table 4 gives the published mid-parent correlations for nine studies. They vary from one exceeding 0·7 (estimated from Cattell & Willson) to 0·35 (Freeman *et al.*). The majority fall a long way short of those predicted, and indicate a large source of variance in parent–offspring correlations unaccounted for by mid-parent IQ. Average weighted mid-parent–offspring correlation: 0·57; unweighted, 0·51.

Correlation and family variance contribution

Two statistical techniques have been used in the investigation of parent–offspring resemblance: correlation and analysis of variance. Strictly speaking, only the former can provide valid data; the latter, buttressed by a genetic model, has been used for this purpose (Eaves, 1973). If the model were correct, the two techniques should provide equivalent results; however, there is reason to suppose that there are important disparities.

The total variance in a set of families consists of within-family variance, between-family variance related to mid-parent IQ and between-family variance unrelated to this. A mid-parent–offspring correlation coefficient only incorporates the between-family IQ related variance as the effect, leaving the rest as error. However, analysis of offspring variance by family groups includes both IQ related and unrelated between-family variance as effects. Thus, the latter is wider than the former. Data were available from two studies in which a direct comparison could be made: Outhit's published results and Alström's unpublished data, which he kindly sent us.

In each case, we have analysed the data to establish whether a significant degree of between-family variance unrelated to mid-parent IQ was present, using a covariance procedure. This removed the between-family correlated variance from the

Table 7. *Evidence for between-family effects unrelated to parent IQ (decimals removed)*

Outhit: mid-parent–offspring correlation 0·665

	Analysis of variance				Analysis of covariance			
	S.S.	d.f.	Var. est.	F	S.S.	d.f.	Var. est.	F
Between-families	48442	50	969	—	16805	50	336	—
Within-families	23100	206	112	8·64	23100	205	113	2·98
Total	71542	256	279	—	39904	255	156	$P < 0·001$

Alström: mid-parent–offspring correlation 0·599

	Analysis of variance				Analysis of covariance			
	S.S.	d.f.	Var. est.	F	S.S.	d.f.	Var. est.	F
Between-families	8714	34	256	—	4659	34	137	—
Within-families	6322	85	74	3·45	6322	84	75	1·82
Total	15036	119	126	—	10981	118	93	$P < 0·025$

total, making it possible to test whether the remaining between-family variance significantly exceeded that within families. The results are shown in Table 7.

It appears that in both studies there is a significant amount of between-family variance unrelated to mid-parent IQ. From this, one must conclude: (*a*) that analysis of variance should not be used as a substitute for correlation to determine the degree of parent–child resemblance, but does provide a useful complement, and (*b*) that there are variables other than parental IQ associated with families, influencing offspring IQ. Test error, of course, does not account for this result since it features in both within- and between-family variance.

Regression

There are two ways of assessing parent–offspring regression which are, theoretically, equivalent. The first involves a comparison of the average deviation of the parents from the test *standardization mean* with the average deviation of the offspring from the same mean. The second uses the parent's deviation from the *parent group mean* and the offspring's deviation from the *offspring group mean*. However, in practice, there are a number of factors which may produce disparities between these two estimates.

The first method will obviously reflect standardization errors, while the second method will not. Secular trends, reflected in intergenerational gains or losses, will operate in the same way as standardization errors, but in this case the choice of method will depend upon one's purpose. Finally, since effects have been found which are unrelated to natural parent intelligence, the way in which these are distributed may modify regression. The possibility must be entertained that the selection procedures used to obtain parent–offspring samples may operate in such a way that they are not truly representative of the total population with respect to these variables. An obvious example is afforded by foster studies.

Table 8. *Group mean regression in non-IQ deviant groups*

Study	Group mean regression (%)			
	Single-parent offspring	Mid-parent–offspring	Single-parent–offspring (SES)	Mid-parent–offspring (SES)
Cattell & Willson	26	—	—	—
Outhit	47	38	—	25 29 (corrected)
Conrad & Jones	49	—	—	—
Higgins *et al.*	54	—	—	—
Alstrom	56	37	—	—
Burks	65	53	—	—
Waller	67	—	49	—
Leahy	—	30	—	20 (indices, not %)
Jones	42	27	—	—

Table 9. *Regression in deviant IQ groups*

Study	Standardization mean regression			Group mean regression	
	Parent IQ	Offspring IQ	Regression (%)	%	Correlation
Polansky *et al.*	78M	94M	73	—	0·35
Snygg	78	95	78	89	0·13
Skodak & Skeels I	86	116	210	97	0·04
II	86	112	184	78	0·28
III	86	115	203	70	0·35
IVa	86	106	144	63	0·38
IVb	86	116	212	59	0·42
Terman/Oden	152	133	36	91	0·08
As adults	134E	133	4	—	—

Notes: M, median; E, estimate.

Group mean regression

(*a*) *Single-parent–offspring regression.* The amount of regression found in each study is shown in Table 8 expressed as a percentage, and in the main calculated from correlation and s.d. data; deviant IQ groups are dealt with separately. The unweighted average is 52 per cent, and the results predictably reflect the same influences as the single-parent–offspring correlations: regression decreases as assortative mating increases, and is higher in 'deviant' groups (see Table 9).

(*b*) *Mid-parent–offspring regression.* Neither the genetic, nor the particular environmental model outlined predicts any regression from mid-parent IQ. It was possible to calculate the regression which actually occurs in the case of five studies: Outhit, 38 per cent; Alström, 37 per cent; Burks, 53 per cent; Jones, 27 per cent; average 39 per cent. Thus, regression from mid-parent IQ appears to be substantial.

(*c*) *Regression from deviant groups.* A summary of the evidence is given in Table 9. There are three studies involving low IQ parent groups, by Polansky *et al.*, Snygg and Skodak & Skeels, the last two being foster studies and thus involving changes of environment unrelated to natural parent IQ.

Standardization mean regression and its relationship to group mean regression

If there are errors involved in the standardization means, the effect of these errors on regression estimates will be greatest for studies involving groups near the population mean. The results thus obtained in fact yield estimates ranging from −200 to +700 per cent. This suggests that standardization mean regression can only be sensibly assessed in groups that deviate substantially from the population mean. Data for the deviant groups are displayed in Table 9 and show little consistency. The sample in Gibson's study (Cambridge scientists) was in effect selected for high IQ and the figure shown denotes the *offspring–parent* regression.

As noted above, of the three studies involving low IQ mothers, in the case of two, the children were raised by foster parents, and thus in environments which were unrelated to natural parents' IQ. The comparability of the results of Polansky *et al.* and Snygg may seem surprising, since one is a foster study and the other is not. There are, however, two complications: first, the children in Polansky's study were in day care and it is possible that the effect of this was similar to fostering, and secondly, it is also possible that while the offspring are similar at their present low age, a difference may emerge as they get older. Other studies (Speer, 1940; Heber *et al.* 1968) have reported progressive IQ declines in children reared by low IQ natural mothers.

The massive degree of regression (past the mean) in Skodak & Skeels' study must reflect considerable environmental influence, there being no reason to suppose that the genetic contribution of the fathers was abnormally high. Although the fathers' IQs were unknown, they were of similar SES and educational level to the mothers. In view of this systematic bias within the group, the disparities between the standardization mean and group mean regression are hardly surprising.

The only study of a deviant group appearing free of problems related to unusual rearing, is that by Terman and Oden, which thus would seem to render it suitable for comparing standardization mean and group mean regression. Taken at face value this study shows that the IQ falls from 152 in the parents to 133 in the offspring, representing 36 per cent regression to the mean, and the result was obtained with very large *n*s. However, there are two possible errors in this result.

The quoted parent mean could be spuriously high due to test unreliability, because the group had been selected for high IQ largely on the basis of a single testing. The size of this error is a matter of conjecture since test reliability for such extreme IQs has not been established, but if one uses the quoted reliability for 'normally' high IQs (i.e. 0·9), then the expected unreliability error would be about 6 points. In addition, since most of the parents were tested on the 1916 Binet and the offspring presumably on the 1937, there are possibly errors arising from discrepancies between the two tests. Such evidence as is available suggests that, as age increases, the 1937 Binet provides higher IQs (Skodak & Skeels, 1949; Freeman *et al.* 1928).

It is unlikely that this degree of regression from the parents' IQs, measured as children, will also occur from their adult status. Longitudinal data on a sample of 72 from the Terman gifted group showed a drop from 150 to 141 from 6 to 12 years of age – less of a decrement than would be expected from our longitudinal data. Seven hundred 13-year-olds from the same sample showed a correlation of 0·66 with the

Concept Mastery Test at age 34. This result corresponds to a drop to 134, and is a little lower than our longitudinal prediction. Using our data and the mid-range age for the Terman group, the predicted adult IQ would be 137. Using the performance on the Concept Mastery Test, McNemar (1947) estimated the IQ of the gifted group as 134 in adulthood. If these two separate lines of evidence are accepted as reasonable estimates of the adult IQs of this group, then near zero parent–offspring regression will have occurred.

To add to these two regression estimates, there is a third estimate – the group mean regression, which for this study is 90 per cent. This abnormally high figure indicates that there is a high degree of within-sample variation unrelated to parent IQ score. Even with random mating, the spouses could not produce more than 50 per cent and, in fact, the estimated spouses' mean IQ on the CMT was 125. We estimate the parent–parent correlation at 0·45 (based on parent IQ variances and the standard error of the difference between couples on the CMT). From the descriptions of the gifted group there is nothing to suggest abnormal within-group variability unrelated to intelligence, yet within this group an adequate degree of variation of intelligence in the parents (s.d. 14·7) does not transmit itself lawfully to the offspring, who, to make matters more puzzling, show an abnormally high degree of variability, significantly greater than that for the parents. This analysis of the complexities within the Terman and Oden study, together with the various estimates of regression (0, 33, 90 per cent) does not accord with the case advanced by Jensen (1973) and is difficult to interpret.

The effect of SES grouping on IQ regression

Three studies provided data which permitted comparisons of parent–offspring regression by IQ alone, by contrast with IQs resulting from SES grouping. If SES had no effect independent of parent IQ on offspring IQ, these two methods of ordering the parents should produce equivalent amounts of regression.

Outhit gives data for mid-parent and offspring IQ by social class which, when put into a regression equation with appropriate *n*s, shows 25 per cent regression by SES. A complication in the Outhit sample was the higher mean of the Alpha-tested children and their preponderance in the higher social classes. Correcting for this increased the figure to 29 per cent (the adjustment took into account the mean difference of the Alpha and Binet children, their proportions in each social class, and the mean differences of their parents, less regression), but even so this is less than regression by mid-parent IQ alone (38 per cent).

A similar analysis of Waller's data gave 49 per cent regression by SES and 67 per cent by IQ alone for single-parent–offspring. For Leahy, parent IQs were not available, but indices based on Otis raw scores gave a figure of 20 by SES and 29 by IQ alone for mid-parent data. Thus, consistently across three studies we found evidence for an effect of SES on child IQ, independent of parent IQ.

Variance

This has two aspects: (*a*) the overall population variance and (*b*) the relative contributions of different subgroups to this total.

Overall variance

The genetic model predicts increasing population variance from parents to offspring with non-random mating whereas the environmental model, in the absence of moderator variables, predicts a decrease. The majority of studies provide data on s.d.s (see Table 4) and with the exception of Oden's, no study shows any appreciable increase in variance from parents to offspring; if anything, the trend is slightly in the reverse direction. Small differences in s.d.s, however, must be interpreted with caution, for tests do not always have exactly the same s.d.s at different ages. Further increments in variance will result from mean differences at various ages when such groups are mixed, particularly in the parent groups.

Subgroup contributions

(1) *Parent deviation from the mean.* Three studies provide data about the effect of the parents' deviation from the mean on the variability of the offspring: Conrad & Jones group the parents into categories according to IQ and Leahy according to SES and thus also by mean IQ. Both of these studies show an essentially uniform child variance irrespective of parent IQ. For Outhit, we correlated mid-parent deviation with the variability of the first four children in each family and obtained a non-significant slightly negative correlation (– 0·2). Overall, these results are inconsistent with the genetic model but fit an 'in ignorance' environmental one.

(2) *Magnitude of IQ difference between husband and wife.* Cattell & Willson report a non-significant positive relationship, whereas both Jones and Outhit report slight negative trends.

DISCUSSION

Taken in conjunction, the results presented in Table 4, and the summaries of the studies, indicate a considerable diversity of outcome and heterogeneity of populations sampled. There are substantial differences in degree of assortative mating, and this has been shown materially to affect variation in single-parent–offspring correlations. In addition, the abnormally low correlations in two of the deviant groups (Oden and Snygg) and, within non-deviant samples, the unusually high and low correlations of Cattell & Willson and Guttman, respectively, suggest within-study sampling variations. It is not possible to conclude with any degree of confidence that one investigation is more representative of a hypothetical 'general population' than is another, and indeed it seems sensible to question whether such a concept has any useful meaning with respect to genetic and environmental components in IQ variation.

In addition to differences in the way variables are distributed within studies, anomalies in the regression data suggest the existence of further overall differences between studies. Thus Burks' sample appears to show 65 per cent within-group regression to a mean of 115, whereas Waller's sample shows virtually the same within-group regression (67 per cent) to a mean of 103.

Without knowledge of the way relevant variables, other than parent intelligence, are distributed in different populations, generalizations, even from large bodies of data such as we have collected, cannot be made with any degree of confidence.

The correlation results indicate that the degree to which offspring IQ is attributable to the IQs of their parents is much less than that predicted on the basis of the genetic and environmental models, individually or combined. The mid-parent–offspring correlations (weighted average 0·57, unweighted 0·51) leave something like 70–75 per cent of the offspring variance not *directly* related to parent IQ. Even when allowance has been made for variance *indirectly* related to parent IQ (either due to random gene segregation or to 'capricious' or 'moderator' environmental variables) one is still left with 35–45 per cent unexplained.

Our analysis of Outhit's and Alström's data shows evidence for effects associated with families, yet unrelated to mid-parent IQ, and the regression data suggest that at least part of this effect is related to SES, again independent of parent IQ.

The possible contribution of test 'error' to the remainder is difficult to assess; it will be recalled that estimates were made of different types, and these may be present to varying degrees according to the design of the investigation. In particular, where studies varied with respect to age and test, it would be reasonable to expect a higher correlation between older children and parents on the same test than between the same parents and younger children, especially when the latter were assessed on a different scale.

Two studies (Outhit and Conrad & Jones), involving closely similar testing procedures, and containing both age and test changes, provide evidence on this point. In each, the children of young and intermediate ages were given the 1916 Binet, and the Army Alpha was used for the older offspring and parents. We estimated potential differences in correlation within these two investigations of up to 30 per cent. In fact, for Outhit, the older offspring–parent correlation was 0·65, and for the younger, 0·61; for Conrad & Jones both were 0·49. Manifestly the overall differences which exist between these samples are not due to variations in age and test. In Kagan & Moss' study, correlations span age 3–10 years, a range over which a difference in these of 50 per cent is possible; yet the results show uniformity. The evidence suggests that information from test equivalence and longitudinal stability may give little guidance for predictions outside these contexts.

Even for the older offspring in the studies discussed above, there was on average a large age disparity between parents and their children. It is possible that this in itself was a source of significant 'error', which further changes of age failed to increase. Of all the investigations, only one involved no age disparity: McCall compared parents tested as children with their offspring at exactly the same ages. The average correlation (0·32) is certainly no higher than those recorded by others, but it must be noted that this was a small select sample.

There remains, of course, the question of how far reliability and standardization errors may contribute to the unaccounted for variance: our crude estimates put these at between 10 and 20 per cent. Compensatory inflation of the mid-parent–offspring correlations to the extent suggested by our upper estimate, together with the corresponding rise in 'capriciousness' would explain all the variance within most studies. However, accounting for all the residual variance in this way ignores its differential distribution within and between families demonstrated in Outhit's and Alström's data. Hence, for these studies at least, such an estimate is obviously too large. Test error, it appears, remains something of an enigma.

On the whole, the data presented in this paper do not permit an assessment of how and to what extent parent intelligence operates environmentally to produce individual differences in IQ score. Certain conclusions can, however, legitimately be drawn. It will be recalled that one entirely unambiguous result was the lack of differential correlations by sex of parent or sex of offspring. It is clear from this that there is no *systematic* tendency for children of one sex to 'identify' intellectually with a parent of the same or opposite sex. Nor were Outhit or Jones able to demonstrate a relationship with the parent of higher or lower intelligence within a family. Assuming that mothers spend more time with their children than do fathers, one might expect a higher correlation with maternal than paternal IQ. Once again the results point to a negative conclusion.

The two foster studies yield lower natural parent–offspring correlations than do those in which children were reared by their own parents, but, as noted, this result is complicated by other factors. The Skodak & Skeels study does, however, indicate a clear environmental effect, which may or may not be related to the IQ of the foster parents, for whom no test data were available.

While we were unable to find, in this group of studies, any single piece of evidence unambiguously demonstrating an environmental effect directly due to parent intelligence, there is one showing a genetic effect, namely the set of natural mother–offspring correlations for ages over two, shown by Skodak & Skeels. However, whether these correlations denote a high or low genetic contribution cannot be assessed, since we do not know what the notional 'correct' parent–offspring correlation should be. In particular, while assortative mating is present in this sample, its degree is unknown; nor is it clear to what extent the children's rearing conditions were homogeneous.

However, good evidence for genetic transmission in one study does not necessitate acceptance of the genetic model, as currently conceived, which, as has been stated earlier, makes a number of precise predictions with respect to parent–offspring resemblances. For children reared with their parents, these predictions were in the main not clearly distinguishable from the environmental ones; however, the polygenic model does entail certain consequences with respect to variance, for which there appears to be no environmental counterpart.

The additive polygenic model predicts that, with assortative mating (which for intelligence is appreciable), the offspring variance should exceed the parent variance. However, only Burt (1955), with estimated parent IQs, was able to fulfil this promise, with an average population, and, as already stated, it is not possible to accept evidence from this source. Furthermore, if assortative mating persists over a number of generations, the trait distribution will become bimodal, and even dominance can only slow down the process. There is no reason to suppose that assortative mating for intelligence is of recent origin, and bimodality should thus have been found in the distribution of IQ scores for children, if not for adults. Test standardization procedures for adults could possibly obscure bimodality, but with children, since the test items are selected on the basis of performance differences *between* ages, bimodality *within* an age group cannot be masked. We have found no report of any such tendency.

The second variance prediction is that offspring variance will be greatest when mid-parent intelligence is average, and will reduce to zero for extreme high or low mid-parent IQ. The addition of other variables (whether environmental or non-IQ related genetic) would reduce, but not eliminate this effect. Three studies reporting variance according to parental level (Conrad & Jones, Leahy and Outhit) found uniformity across the IQ distribution.

The failure of these two genetic predictions is open to one or other of the following interpretations: that a suitable genetic model has been tested, in which case its failures point to an environmental interpretation; or that the particular genetic model is inappropriate and requires modification.

With regard to the latter possibility, geneticists have recorded a number of instances of failures of this model to account for the outcome of breeding studies. The expected reduction in phenotypic variability as selection proceeds is not observed. Admittedly, not all these experiments are suitable for such an analysis, on account of ceiling effects in the measurements used, but those we found which are adequate, fail to show the reduction (De Fries *et al.* 1970; Robertson, 1955).

Hirsch (1962) reports some failures of another prediction, namely that crossing opposite selected strains should produce reduced variance in the F_1 generation, compared with a control group, followed by an increase in F_2. An even more surprising result was reported by Robertson (1955). Two separate strains of *Drosophila* were selectively bred for thorax length until improvement ceased, at which point it might be assumed that further increments were impossible. On crossing the strains, however, further improvement in response to selection over a number of generations was obtained.

A further problem with the present genetic model is its failure to account for the initial appearance of heterosis (hybrid vigour) and its subsequent decline after the F_1 generation (Jones, 1924).

A common explanation for the unexpectedly high variability in the progeny of (non-human) homozygous parents is their increased susceptibility to the environment. However, the only author who cites direct evidence of relative phenotypic variance in response to known environmental changes is Lerner (1954), who analysed data presented by Heuts (1948) and Harvey (1939). Heuts measured the longevity of *Drosophila* in three temperature conditions, and Lerner's analysis showed less change in longevity across varying temperatures for hybrids than for homozygotes. Unfortunately he mistook weeks for days in part of the data, and use of the correct units give results which fail to support Lerner's conclusions. Harvey's data on corn give only incomplete support. It appears, therefore, that there are a number of findings in the field of experimental genetics for which it is not clear whether a genetic or environmental explanation is required.

Our finding of between-family variance unrelated to parent intelligence emphasizes that narrow and wide heritability are separate issues for human intelligence. Furthermore our review, which only concerns part of the evidence on narrow heritability, leads us to conclude that neither our understanding of the mechanisms involved in transmission, nor our knowledge about the origin and nature of sample differences, is adequate to make generalizations concerning parent–offspring resemblances with any degree of confidence.

The authors wish to express their gratitude to Dr Marion Outhit and Dr Marie Skodak for publishing their raw data; to Professor C. H. Alström for making his available; to Dr Marie Oden for data relating to the Terman Gifted Group; to Mr Arthur Royse and Drs D. C. Kendrick and D. Sewell for their valuable comments; to Mr John Devine and his staff for help in using the computer; and to Professor A. D. B. Clarke for his encouragement and advice throughout the research project. We are most grateful to four referees for their helpful evaluations.

References

Adams, M. S. & Neel, J. V. (1967). Children of incest. *Pediatrics* **40**, 55–62.

Alström, C. H. (1961). A study of inheritance of human intelligence. *Acta psychol. neurol. Scand.* **36**, 175–202.

Bayley, N. (1949). Consistency and variability in the growth of intelligence from birth to eighteen years. *J. genet. Psychol.* **75**, 165–196.

Bayley, N. (1957). Data on the growth of intelligence between 16 and 21 years as measured by the Wechsler–Bellevue Scale. *J. genet. Psychol.* **90**, 3–15.

Bloom, B. S. (1964). *Stability and Change in Human Characteristics.* New York: Wiley.

Bradway, K. P. & Thompson, C. W. (1962). Intelligence at adulthood: a twenty-five year follow-up. *J. educ. Psychol.* **53**, 1–14.

Brandon, M. W. G. (1957). The intellectual and social status of children of mental defectives. *J. ment. Sci.* **103**, 710–738.

Burks, B. S. (1928). The relative influence of nature and nurture upon mental development; a comparative study of foster parent–foster child resemblance and true parent–true child resemblance. *Yearb. Nat. Soc. Stud. Educ.* **27** (1), 219–316.

Burt, C. (1943). Ability and income. *Br. J. educ. Psychol.* **13**, 83–98.

Burt, C. (1955). Evidence for the concept of intelligence. *Br. J. educ. Psychol.* **25**, 158–177.

Burt, C. (1961). Intelligence and social mobility. *Br. J. stat. Psychol.* **14**, 3–24.

Cattell, R. B. & Willson, J. L. (1938). Contributions concerning mental inheritance. *Br. J. educ. Psychol.* **8**, 129–149.

Clarke, A. M. & Clarke, A. D. B. (1974). Genetic–environmental interactions in cognitive development. In A. M. Clarke & A. D. B. Clarke (eds), *Mental Deficiency: the Changing Outlook*, 3rd ed. London: Methuen.

Conrad, H. S. & Jones, H. E. (1940). A second study of familial resemblance in intelligence: environmental and genetic implications of parent–child and sibling correlations in the total sample. *Yearb. Nat. Soc. Stud. Educ.* **39** (2), 97–141.

Dearborn, W. F. & Rothney, J. (1941). *Predicting the Child's Development.* Cambridge, Mass.: Sci-Art Publications.

De Fries, J. C., Wilson, J. R. & McClearn, G. E. (1970). Open-field behaviour in mice: selection response and situational generality. *Behav. Gen.* **1**, 195–211.

Doppelt, J. E. & Wallace, W. L. (1955). Standardization of the WAIS for older persons. *J. abnorm. soc. Psychol.* **51**, 312–330.

Eaves, L. J. (1973). Assortative mating and intelligence: an analysis of pedigree data. *Heredity* **30**, 199–210.

Ebert, E. & Simmons, K. (1943). The Brush Foundation study of child growth and development. I. Psychometric tests. *Monogr. Soc. Res. Child Devel.* **8**, No. 2, 1–113.

Eckland, B. K. (1967). Genetics and sociology: a reconsideration. *Am. sociol. Rev.* **32**, 173–194.

Erlenmeyer-Kimling, L. & Jarvik, L. F. (1963). Genetics and intelligence: a review. *Science* **142**, 1477–1479.

Eysenck, H. J. (1973). *The Inequality of Man.* London: Temple Smith.

Foulds, G. A. & Raven, J. C. (1948). Normal changes in the mental abilities of adults as age advances. *J. ment. Sci.* **94**, 133–142.

Freeman, F. N., Holzinger, K. J. & Mitchell, B. C. (1928). The influence of environment on the intelligence, school achievement and conduct of foster children. *Yearb. Nat. Soc. Stud. Educ.* **27** (1), 219–316.

Gibson, J. B. (1970). Biological aspects of a high socio-economic group. I. IQ, education and social mobility. *J. biosoc. Sci.* **2**, 1–16.

Guttman, R. (1974). Genetic analysis of analytical spatial ability: Raven's Progress Matrices. *Behav. Gen.* **4**, 273–283.

Halperin, S. L. (1945). A clinico-genetical study of mental defect. *Am. J. ment. Defic.* **50**, 8–26.

Harvey, P. H. (1939). Hereditary variation in plant nutrition. *Genetics* **24**, 437–461.

Heber, R., Dever, R. & Conry, J. (1968). The influence of environmental and genetic variables on intellectual development. In H. J. Prehm, L. A. Hamerlynck & J. E. Crosson (eds), *Behavioural Research in Mental Retardation*. Eugene, Oregon: University of Oregon Press, pp. 1–23.

Herrnstein, R. J. (1973). *IQ in the Meritocracy*. London: Allen Lane.

Heuts, M. J. (1948). Adaptive properties of carriers of certain gene arrangements in *Drosophila pseudoobscura*. *Heredity* **2**, 63–75.

Higgins, J. V., Reed, E. W. & Reed, S. C. (1962). Intelligence and family size: a paradox resolved. *Eugen. Q.* **9**, 84–90.

Hirsch, J. (1962). Individual differences in behaviour and their genetic basis. In E. L. Bliss (ed.), *Roots of Behaviour*. New York: P. Hoeberg.

Hirsch, N. D. M. (1930). An experimental study upon three hundred school children over a six year period. *Genet. psychol. Monogr.* **7**, No. 6, 493–546.

Honzik, M. P. (1957). Developmental studies of parent–child resemblance in intelligence. *Child Devel.* **28**, 215–228.

Honzik, M. P. (1972). Intellectual abilities at age 40 years in relation to the early family environment. In F. J. Monks, W. W. Hartup & J. de Wit (eds), *Determinants of Behavioural Development*. New York: Academic Press.

Honzik, M. P., McFarlane, J. W. & Allen, L. (1938). The stability of mental test performance between two and eighteen years. *J. exp. Educ.* **17**, 309–324.

Jencks, C. (1973). *Inequality*. London: Allen Lane.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harv. Educ. Rev.* **39**, 1–123.

Jensen, A. R. (1973). *Educability and Group Differences*. London: Methuen.

Jensen, A. R. (1974*a*). Cumulative deficit: a testable hypothesis? *Dev. Psychol.* **10**, 996–1019.

Jensen, A. R. (1974*b*). Kinship correlations reported by Sir Cyril Burt. *Behav. Gen.* **4**, 1–28.

Jones, D. F. (1924). The attainment of homozygosity in inbred strains of maize. *Genetics* **9**, 405–418.

Jones, H. E. (1928). A first study of parent–child resemblance in intelligence. *Yearb. Nat. Soc. Stud. Educ.* **27** (1), 61–72.

Kagan, J. & Moss, H. A. (1959). Parental correlates of child's IQ and height: a cross-validation of the Berkeley growth study results. *Child Dev.* **30**, 325–332.

Kamin, L. (1973). Heredity, intelligence, politics, and psychology. Invited address to the Eastern Psychol. Assoc. Meetings, May 1973. Subsequently published in C. J. Karier (ed.), *Shaping the Educational State: 1900–1970*. New York: Free Press, 1975.

Kangas, J. & Bradway, K. P. (1971). Intelligence at middle age: a 38-year follow-up. *Dev. Psychol.* **5**, 333–337.

Leahy, A. M. (1935). Nature–nurture and intelligence. *Genet. Psychol. Monogr.* **17** (4), 241–305.

Lerner, I. M. (1954). *Genetic Homeostasis*. Edinburgh: Oliver & Boyd.

McCall, R. B. (1970). Intelligence quotient pattern over age: comparisons among siblings and parent–child pairs. *Science* **170**, 644–648.

McNemar, Q. (1942). *The Revision of the Stanford–Binet Scale: an Analysis of the Standardization Data*. Boston: Houghton Mifflin.

McNemar, Q. (1947). Intellectual status of the gifted subjects as adults. In L. M. Terman & M. H. Oden (eds), *Genetic Studies of Genius*, IV. Stanford University Press.

Outhit, M. C. (1933). A study of the resemblance of parents and children in general intelligence. *Archs Psychol.* **23**, 1–60.

Penrose, L. S. (1933). A study of the inheritance of intelligence. *Br. J. Psychol.* **24**, 1–19.

Pinneau, S. R. (1961). *Changes in Intelligence from Infancy to Maturity*. Boston: Houghton Mifflin.

Polansky, N. A., Borgman, R. D. & De Saix, C. (1969). Mental organization and maternal adequacy in rural Appalachia. *Am. J. Orthopsychiat.* **39**, 246–247.

Reed, E. W. & Reed, S. C. (1965). *Mental Retardation: A Family Study*. Philadelphia: Saunders.

Robertson, F. W. (1955). Selection response and the properties of genetic variation. *Coldspring Harb. Symp. Quant. Biol.* **20**, 166–177.

Schaie, K. W. & Strother, C. R. (1968). A cross-sequential study of age changes in cognitive behaviour. *Psychol. Bull.* **70**, 671–680.

Skodak, M. & Skeels, H. M. (1949). A final follow-up study of 100 adopted children. *J. genet. Psychol.* **75**, 85–125.

Snygg, D. (1938). The relation between the intelligence of mothers and of their children living in foster homes. *J. genet. Psychol.* **52**, 401–406.

Sontag, L. W., Baker, C. T. & Nelson, V. L. (1958). Mental growth and personality development: a longitudinal study. *Monogr. Soc. Res. Child Devel.* **23**, No. 68, 1–143.

Speer, G. S. (1940). The mental development of children of feeble-minded and normal mothers. *Yearb. Nat. Soc. Stud. Educ.* **39** (2), 309–314.

Stanton, R. G. (1946). Filial and fraternal correlations in successive generations. *Ann. Eugen.* **13**, 18–24.

Waller, J. H. (1971). Achievement and social mobility: relationships among IQ score, education and occupation in two generations. *Soc. Biol.* **18**, 252–259.

Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence*, 4th ed. Baltimore: Williams & Wilkins.

Willoughby, R. R. (1928). Family similarities in mental test abilities. *Yearb. Nat. Soc. Stud. Educ.* **27**, No. 1, 55–59.

(*Manuscript received* 21 *April* 1975; *revised manuscript received* 30 *July* 1975)

Br. J. Psychol. (1976), 76, 2, *pp.* 275–296
Printed in Great Britain

BOOK REVIEWS

The Illusion of Attitude Change: Towards a Response Contagion Theory of Persuasion. By J. M. Nuttin, Snr. London: Academic Press. 1975. Pp. xii + 236. £6.40.

This is a remarkable book. Its impact on the study of attitude change will be considerable.

Attitude change has been in the doldrums for some time. Much information is available from the study of the effect of persuasive messages on opinion change and for some time it has seemed unexciting to discover more. Now, according to Nuttin, our knowledge is illusory. We have not, apparently, been studying attitude change at all when examining the impact of persuasive messages (whether from an external source, or from oneself). Changes occur in the long term *only* if they have been measured once before, shortly after exposure to the persuasive message, and such 'changes' as do occur are limited to statements formulated in the message, and do not have any impact on closely related expressions of opinion. Moreover, it seems that cognitive explanations of changes in terms of principles of balance, consonance or some other principle of organization, may be effectively attacked.

The attack on complex cognitive theories of attitude change is exciting to read. Quaint English cannot mask the author's clarity of thought and scientific zeal. Evidence is weighed and inquiries pursued with the skill of a good detective writer. Some experiments are designed to throw light on old controversies, and merely set the scene for more interesting later experiments specifically designed to test Nuttin's own theory. Fortunately, a bitter personal controversy with Rosenberg is relegated to an Appendix.

Nuttin asserts that responses to questionnaire items are as subject to situational control as in any other species of overt behaviour. They are 'mere evaluative responses whose determinants can be studied without reference to the underlying attitude construct' (p. 191). He is led to adopt this view both because of the failure to demonstrate long-term effects and because the results of a series of ingenious and well-controlled experiments suggest that the explanation of some classic dissonance effects may be much less complex than previously envisaged.

The inverse relationship between size of reward and opinion change as a result of 'counter-attitudinal advocacy' in situations of 'forced compliance' is the most notorious of the classic dissonance effects. Is it reliable? In a series of carefully constructed replications Nuttin shows that neither the dissonance nor Rosenberg's effective–cognitive consistency theory give plausible explanations of the effects obtained. For example, he employed two crucial control conditions: an 'advocacy for free' condition and a pro-attitudinal advocacy condition. If anything, advocacy of counter-attitudinal material with no payment produced less change than paid advocacy, and pro-attitudinal advocacy produced results that were by and large indistinguishable from those of counter-attitudinal conditions. In another experiment it was shown most realistically that anonymity does not lessen the degree of the change. Whilst certain effects remained consistent with dissonance theory, their explanation in terms of dissonance theory was severely questioned by the inconsistencies I have listed. For example, in one set of comparisons 'free advocacy' (both pro- and counter-) produced about the same degree of change as a large reward for counter-attitudinal advocacy. This is difficult to reconcile with dissonance theory. However, a low reward and a 'relative deprivation' condition produced considerably more change than the high reward condition. This result is, indeed, consistent with dissonance theory, but the plausibility of dissonance theory is already in doubt. Nuttin's way out is to suggest that the low reward and relative deprivation conditions had something else in common that brought about change in only those two conditions: their unexpected socially embarrassing character. A few pence for making a nationwide television counter-attitudinal plea is *more* embarrassing than no reward. In the relative deprivation condition the subject was given no money but was informed that previous subjects had been paid—as equally socially embarrassing, and inappropriate a piece of information as a derisory reward.

The ability of embarrassing circumstances to produce attitude change in counter-attitudinal advocacy is illustrated in a series of ingenious, elaborate and very carefully constructed experiments, the ethics of which in some cases were dubious. In one telling experiment, the supposed reward of a faked pass in an important university examination was unanimously preferred (by all students 'fortunate' enough to be in that experimental condition) to a high monetary reward. Rather than this magnificent reward eliciting less change (as would be predicted by dissonance theory) it produced an enormous change in the direction of the counter-attitudinal plea. In another experiment the close presence of an attractive female (experimenter) don wearing 'hot pants' whilst directing the experiment in the highly conservative and male-dominated University of Leuven was sufficient to induce massive change in accordance with the required counter-attitudinal (or pro-attitudinal) plea. Care was taken in this and other experiments using different 'embarrassing' events to rule out explanations in terms of social desirability.

How does Nuttin explain the effectiveness of embarrassment in inducing subjects to 'assimilate' their 'evaluation responses' to those of the prior message? It is, says Nuttin, the result of *response contagion*, and is a function of three factors: '(1) the nature of the prior responses; (2) the degree of perturbation or arousal; (3) the temporal continuity between prior and subsequent responses'. Principally, 'perturbation' or 'arousal' facilitates the dominant, prior responses of the message, especially when the expression of the subsequent evaluative response follows shortly after the 'message' has been delivered. In a closing chapter, Nuttin attempts to show that the results of a number of well-known attitude-change studies may be given a more or less convincing explanation along these lines.

Nuttin and his experiments are persuasive, but we should remember that there is, as yet, no *direct* physiological evidence for Nuttin's account. He has cast doubt on complex cognitive theories, and has produced results which accord with this theory of response contagion. No doubt future work will test directly the supposed 'perturbing' effects of different experimental manipulations and examine afresh their relationship with subsequent responses.

Despite a characteristically meticulous attention to detail, a purist might wish to find fault with the occasional tendency to combine data from different experiments. However, this is rarely done without good reason. The implications of his work are not immediately apparent. Nuttin suggests 'abandoning the social attitude concept because it is both misleading and useless for predicting, controlling and understanding social behaviour' (p. 216). This exaggerates the importance of his findings. Nuttin has apparently shown that much of the so-called 'attitude change' literature is not dealing with changes in social attitudes, but with a species of behavioural contagion. This does not, in itself, invalidate the social attitude concept, nor, indeed, the truth of dissonance theory as applied to the organization of social attitudes.

According to Nuttin, we should continue our studies of behavioural contagion in this way because it may throw light on important social processes like psychotherapy, advertising, prayer, conversion and crowd behaviour, and because it may help bring general and social psychology into closer relation. Perhaps he is right. Having discovered a wolf in sheep's clothing we should not shoot it dead. Apart from which, this particular wolf has been around since the time of Le Bon, and a lot of us have grown rather fond of him by now, and will be only too glad to celebrate his re-emergence. The experimental study of attitude change is, in comparison, but a tender young lambkin. However, it is a popular pet and not so ready for slaughter as Nuttin suggests.

GEOFFREY M. STEPHENSON

Society, Stress and Disease. Volume 2. *Childhood and Adolescence*. Edited by LENNART LEVI. London: Oxford University Press. 1975. Pp. xxii+551. £22.50.

During June and July 1971 a symposium was held in Stockholm, jointly organized by the University of Uppsala and the World Health Organization. The purpose of the symposium was to consider the role of social factors in preventing or causing disease during the age from conception to adolescence. It was the second symposium in a planned series with the theme 'Society, Stress and Disease', and the 54 papers which formed the basis of the symposium are presented in this second volume. There is a lot about stress, a little about disease, and next to nothing about society.

The first session of the symposium was on the 'Well-adapted Child', and the best contribution is by Lazarus on 'The healthy personality', a review of writing and research with the aim of finding answers to the question as to what constitutes psychological health. Other authors contribute various discourses on the topic: on social hierarchies; on psychosocial factors in growth and development; on adaptations in early marriage, infancy and the preschool period; on culture and inbreeding; on stress, homeostasis and time. Of course there are many answers to this crucial question; Fromm's 'productive personality', Grinker's 'homoclites', and even power behaviour amongst macaque monkeys is offered for serious consideration, together with many other formulations. The list invites contribution and the present writer offers 'rational-altruistic-democratic man' for serious inclusion, although he has never met a person with all of those traits adequately developed, and probably never will. However, idealistic conceptions are important. They remind us that, psychopathologically, smug vapidity or coarse one-upmanship are not indices of health; they merely indicate compromise coping with 'sick' societal arrangements.

The second session had an aweseome title, 'Potentially Pathogenic Psychosocial Stressors, Intervening Variables, Mechanisms and Precursors, in Childhood and Adolescence', and 17 papers are presented which list the industrial society, the mass media, the violent society, the educational system, social unrest, normal childhood, negative social inheritance, parental rejection, paternal and maternal deprivation, adoptive background, physical and mental handicap, and malnutrition as probable psychosocial stressors. Some of these are good methodological studies, others are rather odd contributions. An example of the former is Frankenhausser and Johansson's 'Behaviour and catecholamines in children', which is just about that, and which also gives a fascinating vignette of parental punishment and adrenaline reaction in children. An example of an odd curio is Corson and Corson's 'Education and the adolescent', which is more about 'biological individuality' and 'Pavlovian typology' than it is about education or adolescence. The editor must have had this session in mind when he expressed the prefatory hope that the volume presented 'not only a selection but also a meaningful pattern of facts, theories, hypotheses, speculations, and value judgements concerning the various components of the "psychosocial factors–stress–disease" system and their interrelationship . . .'. Considering the sheer diversity of content, a meaningful pattern is elusive. Only a Bronfennbrenner could distil the essential meaning of this plethora of 'psychosocial' events, although a potent degree of 'proof' would be unlikely.

Given the imprecision of knowledge about stressors it could hardly be expected that 'Psychiatric and Psychosocial Disease Processes', the topic of the third session, would be precisely delineated, and indeed it was not. However, because response patterns are easier to define, to do research on, and to write about, this session produced a number of sound and interesting methodological studies, offset by several goodish reviews of findings in the field. The best contribution is by Anokhina, Zabrodina and Svirinovsky, 'General mechanisms of development of psychopathological reactions under emotional stress—clinical and experimental research': this work does teach us the meaning of stress as a physiochemical event. Other studies are not so epistemically incisive, although increasing our understanding of certain areas: neonatal behaviour as a predictor of behaviour in childhood stress situations; automatic reactivity and behavioural deviation; development of psychopathological reactions under emotional stress; negative social inheritance and biological and non-biological family rearing; ethnic group, immigration, and infant morbidity; the mother and her first born. All the reviews are excellent accounts of health and environment, effects of exposure to stressors in early life upon later behaviour and biology, psychosocial stressors and intellectual deficits. The most excellent contribution, however, is Kagan and Levi's offsetting review of what is actually 'known' about the relationship between psychosocial stimuli and pathological response patterns; they conclude that there is a shortage of real proof.

Given the imprecision of knowledge about stressors, and the lack of proof as to their effect on individual functioning, it follows that knowledge of specific diseases due to psychosocial stress, and their cure and prevention, will also be imprecise and short on proof. Such is the case as demonstrated by the contributions of sessions 4 and 5 of the symposium. Contributions with a psychosocial content loom largest; psychosomatic diseases and antisocial behaviour patterns. Of the former respiratory diseases, gastrointestinal disorders, migraine, eczema, obesity, and other disorders receive for the most part speculative treatment. Of the antisocial phenomena adolescent suicide, drug abuse and the battered-child phenomenon are interestingly written about; the article on drug-taking provides a charming little glossary to educate 'straights'. Interesting writing, however, is not a substitute for sound methodological studies that convincingly trace the association between stress and disease. In the absence of these it can only be expected that proposals as to prevention would be pedestrian, and such is also the case. One exception to this is the article by Petros-Barvazian of the World Health Organization, namely

'Some priority problems in maternal and child health—psychological aspects and preventive measures'; it is exceptional in that clear preventive measures follow from the research findings obtained from many countries. Much lip-service is paid to the concept of cost/benefit analysis, but on the whole neither cost or benefit could be adequately estimated from such work. Applied research must give rise to clear-cut preventive formulations which can be adequately evaluated otherwise the consumer is entitled to query the waste of public funds; this is a conclusion strongly implied by the only paper in session 6, 'Future research' by Engström.

In the final summing-up it could be argued that this is essentially a pathfinder book and that reconnaissance is notoriously fraught with ambiguity. However, it is more a Noah's Ark collection; it is sufficient to indicate that there are tides of adversity in society and what kind of social animals are likely to be effected by these. On the other hand, little in the way of a course is plotted which outlines likely causes, possible cures and effective prevention. This is because the evidence presented does not reach an acceptable level of proof on the one hand, and also because the frame of reference of the whole endeavour is substantially incomplete, on the other. With respect to the first point, Kagan and Levi in their paper 'Health and environment—psychosocial stimuli: a review' conclude that 'causation of disease by psychosocial stimuli is unproven but at a high level of suspicion'. What demonstrations are necessary to go beyond suspicion to accepted fact? Probably no amount of demonstrations will ever add up to acceptable proof because the universe of discourse is much too limited. As pointed out at the beginning of this review, the volume says next to nothing about society, except that it is something which gives rise to psychosocial stimuli. To reduce society to this for the sake of expediting the usual mechanical model will simply not do nowadays. Alexander Leighton's Stirling County study (1959) shows how far social psychiatry can go in tracing down the noxious effects of society. Issues of personal freedom, of economic freedom, of rationality, of social arrangements, type of family, goals for achieving statuses, means for achieving them all came up for critical review.

J. MCALLISTER

REFERENCE

LEIGHTON, A. H. *et al.* (1959). *My Name is Legion* (volume 1 of the Stirling County Study). New York: Basic Books.

Stressful Life Events: Their Nature and Effects. Edited by B. S. DOHRENWEND and B. P. DOHRENWEND. New York: Wiley. 1974. Pp. xi + 340. £10.80.

The possibility that stressful life events may play a part in the aetiology of various somatic and psychiatric illnesses has radical and extensive implications for the practice of medicine. If life experiences can help cause illness, then conventional medicine with its emphasis on the patient as a mere biochemical machine becomes obsolete in the narrowness of its scope, and demands to be extended to consider the total social and environmental fate of its clients.

With such widespread implications at stake there has probably been a tendency to overestimate the validity of positive findings and to ignore the substantial methodological problems still to be overcome. With such unwarranted extrapolations being performed there is the danger of an over-reaction on the part of the inferentially cautious leading to a premature rejection of the whole debate.

The Dohrenwends' book comes at a critical point in the subject's development. It restricts itself to a cautious, thorough and thoughtful analysis of the evidence to date pointing out the limit of empirical findings which have as yet been produced and the related theoretical and methodological problems in which the research is still entangled. It argues for the ultimate value of the research area but is sceptical about its present production.

The contents of the book are based on a series of papers given at the 1973 Conference on Stressful Life Events. As editors, B. S. Dohrenwend and B. P. Dohrenwend have grouped the papers into five topic areas, arranged them in logical sequence and written an overall introduction, with introductions to each section and a closing summative chapter. This arrangement, together with the papers, allows the issues to unfold in a clear and logical manner, and makes the book attractively readable even to the person with little knowledge of the area.

Part I deals with the relationship of stressful events and physical illness. Hinkle, in a concise review of 20 years of his own empirical investigations, ends with the cautious conclusion that although social change is associated with health changes, its effects cannot be a single function of the change itself, since other crucial variables are the physical and psychological characteristics of the individual which act as a mediating process. The problems of defining and measuring life stress are explored in two papers, by Rahe and by Holmes and Masuda, which offer the schedule of recent experiences as a solution. Mechanic in his paper raises a very basic issue that any measurement of stress must depend on the existence of an adequate conceptual account of the phenomenon.

Part II consists of studies which consider the relationship between particular types of life events and particular types of illness and which employ a common research method, namely the comparing of a patient's group with matched controls. Evidence is reported that various types of life events are predictive at myocardial infarction, adolescent depression, schizophrenia and severity of neurotic symptomatology. What prevents generalizing from these studies is, firstly, the different measures they each employ and, secondly, the fault they share in common which is relying on retrospective reporting of stress experiences. These criticisms are sufficient for one of the authors, Hudgens, to treat the case as not yet proven.

The community studies reported in Part III relying as they do on methods derived from epidemiology overcome some of the problems. In particular they offer the possibility of prospective studies which allow some predictions and can avoid the contamination of dependent and independent variables. In addition they produce a link between life experiences and community variables such as social class.

Parts IV and V consist of an intensive critical analysis of the shortcomings in the literature to the present time. The present consensus scaling of life events as used in the Holmes and Rahe scales is considered by Brown to be both theoretically naïve and dangerously misleading. A typology of life events in terms of gain versus loss; controllable versus uncontrollable, expected versus unexpected, and objective versus subjective is recommended by Dohrenwend. The final paper by the two editors draws together the strands of the debates and argues for a proper sampling of all life events, prospective research designs, and a consideration of all types of illness outcome.

The scope of this book is wide, covering both physical and psychiatric illness, though the emphasis is probably on the latter, given the interests of the editors. It makes intelligible and stimulating reading. My only quibble is the price: £10.80 is rather a lot to pay for research reported in 1973.

PATRICK J. O'DONNELL

The Origins of Fear. Edited by M. R. LEWIS AND L. A. ROSENBLUM. New York: Wiley. 1974. Pp. xiii+284. £9.70.

To those who have occasionally suspected current 'developmental' psychology to be a complex blend of the incomprehensibilities of Piaget and radical politics, this collection of modified conference papers will come as a pleasant surprise. The contributing authors are invariably cheerful, modest and open-minded; they all share sympathies for the experimental method and for British empiricism; and they all expand liberally and intelligibly on their chosen themes.

Less pleasing, and of less surprise, will be the editors' implicit contention that their book's overwhelming concentration upon experimental work with infants is an adequate response to the challenge of the title. Concepts such as heritability and conditioning are spurned, out of preference for the more malleable notion of 'assimilation'; the role of assimilation in determining fearfulness is much pondered, but no reference is made to intelligence; curvilinear relations between arousal and performance are seldom illuminated by reference to the relevant one of psychology's few laws; and the individual differences – some of them 'enormous' – exposed in these studies are never traced back to any origin whatsoever. Some of the authors freely admit that their own studies of fear may have little light to shed on how fear arises and persists in adulthood or even childhood: perhaps for this reason references to such concepts as 'anxiety', 'phobia', 'neuroticism' and 'ego-strength' are virtually non-existent.

But what, then, is to be assimilated by the psychiatrists, paediatricians and educational psychologists to whom the publishers commend these 'authoritative research findings'? There is much rehearsal of experimental evidence supporting the notion that proximity to the mother (or, indeed, the father) reduces fear and increases exploration of novel stimuli: P. K. Smith discovers that object manipulation goes up as automanipulation goes down; and M. Lewis and J. Brooks report 'obtaining the services of a small adult to see whether her approach would elicit the response pattern characteristic of a child or of an adult' – the result being somewhere in between. More seriously, Sroufe shows the role of context in determining whether an infant laughs or cries to a novel stimulus; and suggests that level of arousal alone cannot account for the way in which an infant seems to 'judge' the threat-value of a stimulus. K. Kagan sympathizes with this, and stresses possible temperamental determinants of reactions to discrepancy. Schaffer considers how impulsiveness gives way to response-inhibition in the first year of life; and, like other authors, considers the development of object concepts relevant to understanding the development of fear of strangers.

These are serious and thoughtful pieces; and there are others to be found. But they are at some remove from the lucid argumentation and demonstration to which Bower and Bryant have accustomed serious 'developmental' audiences in the recent past. And they are workaday communications from one laboratory to another rather than papers which could commend the book to a wider audience. Even at this level, Harriett Rheingold's concluding comments on the papers are appropriately cautionary: it is unclear what information of consequence for understanding development is supposed to be gleaned from the customary laboratory procedures employed; 'fear' itself is never unambiguously defined in these papers; measures of inter-observer agreement are scantily reported; and the 'perceptual-cognitive' approach that has replaced most other kinds of psychology in these authors' minds is one that seems to have little predictive power.

CHRIS BRAND

Principles of Psychotherapy. By I. B. WEINER. New York: Wiley. 1975. Pp. xii + 332. £9.80.

In this book. I. B. Weiner sets out to help practitioners of psychotherapy to understand the psychotherapeutic process in the light of his extensive clinical experience and modern psychological research and thus hopefully to become more effective.

Psychotherapy is carefully defined at the outset – a definition which makes clear the eclectic but psychodynamic orientation of the author: 'Psychotherapy is an interpersonal process in which one person communicates to another that he understands him, respects him and wants to help him.' Although this definition is very wide and embraces most respectable methods of psychotherapy, it obviously implies an approach that is not highly relevant to behaviour or drug therapy. Nevertheless, with this proviso, Weiner manages to include discussions on most of the topics that psychotherapists need to consider, as a sample of the contents indicates: the aims of psychotherapy, patient variables and therapist variables that affect the outcome, the interpersonal skills needed by the therapist, the theoretical basis of personality dynamics (neo-Freudian this but not dogmatically so, although it will frighten off 'scientific' psychotherapists no doubt) and the psychotherapeutic process. This is only the beginning. A similar analytic procedure is applied to the treatment, carefully divided into phases and demonstrating what is going on at any given time.

To give reality to these discussions – for although Weiner refers to research wherever possible he is well aware that much clinical research is notable for the elegance of its design rather than its pertinence to psychotherapy – case material is frequently introduced. This clinical evidence is well chosen to support and illustrate his arguments. However, two points must be noted. It lacks the amazing clarity of the cases quoted in Fenichel (1945), which give a profound and memorable insight into psychoanalytic theory even if they seem far-fetched to any but believers. In addition, Weiner's cases, as the objections to Socrates in Platonic dialogues, are *too* convenient. The reader is left feeling, 'All right, but suppose ... then what?' This is an inevitable weakness of trying to support an argument by this type of evidence. From this brief description of the book it is clear that Weiner has covered the material that psychotherapists are concerned with in a way that is obviously relevant to their problems and difficulties. In this respect, the book is most carefully laid out with manifold and accurate subheadings so that any specific point can be easily found.

Given the frankly psychodynamic orientation, Weiner's exposition is clear and any clinical psychologist who could not grasp his points and see their application to his own work should not be in practice. Chapter 10, where the psychotherapeutic relationship is examined, is a good example to take, since few therapists would deny some importance to this aspect of their work. The relationship is here seen on three levels: transference, reality and the working alliance.

Great importance is attached to therapists being able to disentangle these three components of the relationship so that the psychotherapeutic events can be correctly interpreted. Obviously, in these terms, to interpret a reality-based response of a patient as a transference-based response would be fatal. Through case material Weiner shows how the components can be recognized, the aim being to understand the transference relationship. This chapter highlights the strengths and weaknesses of the book as a teaching aid. It is stressed all the time that the relationship of therapist and patient is not like those in the real world. This is the force of the working alliance. Similarly the real relationship is slight because of the transference feelings and the working alliance. Therapists are advised to keep it this way without thereby destroying the Truax canons of empathy, warmth and genuineness. This is fine *given the psychodynamic orientation*. However, this contrasts strongly, for example, with the therapeutic regime advocated by Lomas in *Time and False Experience*, where it is well argued that therapy can give a patient perhaps his first experience of real trusting and genuine relationship and that this is highly beneficial. Again there is no reference to the role-playing aspects of relationships stressed by Berne and his transactional colleagues.

In brief, this is a clear, careful discussion of most aspects of psychotherapy, but limited, to some extent, by the psychodynamic orientation. However, one does not have to agree with this outlook to benefit from the discussion and it is a book that can be recommended to all practitioners, future and present, of psychotherapy. However, it is one-sided and for a full picture other books would need to be studied.

P. KLINE

References

Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: Norton.
Lomas, P. (1973). *True and False Experience*. London: Allen Lane.

The Psychoanalysis of Artistic Vision and Hearing. An Introduction to a Theory of Unconscious Perception. By Anton Ehrenzweig. London: Sheldon Press. 1975. Pp. xxxii+272. £5.95.

Described as a third edition, this book is in fact a re-issue of a work originally published in 1953; it was then reviewed by the late Carolus Oldfield [*Br. J. Psychol.* (1954), **45**, 152]. A second, unaltered, edition appeared in 1965. Ehrenzweig died in 1967. He had been a lecturer in art education at Goldsmith's College, and in the preface to the second edition he expressed disappointment that his book, while welcomed by artists (including Gombrich and Herbert Read), should have been virtually ignored by psychologists. Oldfield appears to have been an exception, inasmuch as he commends Ehrenzweig's courage in tackling problems in perception on which 'official' psychological theories are silent.

Someone once said that psychoanalysis is spelled without a hyphen by its adherents, and by others with. On this criterion Ehrenzweig would appear to belong to the latter category, and while Freudian concepts (and others deriving more specifically from Melanie Klein) do of course provide the core material of the book, it is interesting that the starting point of the argument is to be found in the work not of Freud but of William James. An epigraph on the title-page refers to James's concern for 'the reinstatement of the vague to its proper place in mental life', and what James called the 'Psychologist's fallacy' is a theme that may be said to permeate the whole book. James's original formulation of this 'fallacy', to be found in volume I of the *Principles* (p. 196), refers to 'the confusion of [the psychologist's] own standpoint with that of the mental fact about which he is making his report'. This confusion Ehrenzweig reformulates as 'the observing mind's tendency to perceive only articulate form', a tendency basic both to Gestalt theory and to Freud's observation that 'form experiences coming from lower layers of mind, like our dream visions, tend to be inarticulate' and consequently 'chaotic and difficult to grasp'. Perception of the gestalt, according to Ehrenzweig, is a function of what he calls the 'surface mind'; what lies outside the gestalt is the concern of the 'depth mind' – or of the unconscious, although this latter is a term of which Ehrenzweig makes sparing use. Aesthetic perception, and the creative experience of the artist, are both concerned with the inarticulate, or more properly with the thing, as distinct from the gestalt. 'The work of art', Ehrenzweig writes, 'remains the *Ding an sich*, about the real structure of which' – owing to the limitations of our observing mind – 'we cannot make any definite statements'.

The foregoing attempt to outline Ehrenzweig's position has inevitably led to oversimplification. It is a difficult book to read, let alone to summarize. In particular it is impossible to set out, in short compass, the author's treatment of specifically psychoanalytic themes, but it may be noted that he extends the concept of repression to cover all failure to perceive 'the inarticulate' or whatever is 'outside the gestalt'. Included are the creative artist's 'fingerprints' (Ehrenzweig prefers the term 'handwriting') and stylistic idiosyncrasies in musical performance. It is interesting to note, incidentally, that Ehrenzweig sees music, rather than visual art, as particularly amenable to depth-psychological analysis, since it embodies 'a sharp distinction between articulate and inarticulate form elements'. In this respect his book is in marked contrast to the approach of, say, Erich Neumann in his *Art and the Creative Unconscious* (Routledge & Kegan Paul, 1959), or to Freud's own in his *Leonardo*. What Ehrenzweig presents is essentially a metapsychology, and it is this, no doubt, that Oldfield had in mind when he called attention to the difficulty Ehrenzweig had to face in discussing what orthodox psychological theory is 'framed expressly to exclude from consideration'.

The reader who is prepared to go along with Ehrenzweig will find the book rewarding. One complaint: it is a pity that in the twenty-odd years since it first appeared no one should have thought to provide it with an adequate index. As has already been indicated, it is a book that requires close reading, during which one frequently wants to refer back to something that has been said earlier, and it is disconcerting to find that many of the author's key concepts are totally missing from the index.

BORIS SEMEONOFF

Hemispheric Disconnection and Cerebral Function. Edited by M. KINSBOURNE AND W. LYNN SMITH. Springfield, Ill.: Thomas. 1974. Pp. xvi + 300. $21.50.

This book asks what we may learn from two surgical operations – hemispherectomy and commissurotomy. Both operations are performed as extreme measures and both, if the patient survives, have a less devastating effect on everyday behaviour than might have been predicted. At least one patient has obtained a university diploma after removal of the right hemisphere in early adulthood. Early in the book Aaron Smith introduces an important but neglected paradox of neuropsychology, a paradox to which Kinsbourne returns in a later chapter: 'the function of the whole appears to be less than the sum of its parts'. Focal lesions within a lobe may result in deficits that are not observed after the loss of the entire lobe; and focal lesions within a hemisphere may produce deficits that are not found after hemispherectomy. Whereas a lesion of the left occipital lobe may lead to colour agnosia, left hemispherectomy does not; and whereas right occipital disease may, it is said, result in prosopagnosia, right hemispherectomy does not. Left hemispherectomy in the adult does cause severe aphasia, but Smith reports that a dysarthria present before operation cleared up after left hemispherectomy, and auditory comprehension returned to normal levels. Neither his right, nor his left, hemispherectomy patients showed the word deafness or the total constructional apraxia or the amusia that may be produced by restricted lesions in one hemisphere. To account for these discrepancies, Kinsbourne proposes that areas of cortex subserving specific functions may inhibit other areas that are potentially capable of the same function: a diseased region of cortex may prevent compensatory activity by healthy regions of the contralateral cortex.

Kinsbourne's explanation must be judged within the more general theory of hemispheric interaction that he develops in this book. However, cases of adult hemispherectomy are very few and he himself acknowledges one major difficulty in interpreting the results: the disease is often of long standing and the relatively healthy hemisphere might have taken on compensatory functions many years before the operation while the brain was still plastic. It has long, of course, been known that hemispherectomy can be remarkably mild in its permanent effects if the operation is performed in infancy. In such cases the later level of verbal intelligence appears to be unrelated to the laterality of disease. The later the operation the less complete the compensation. Kohn and Dennis, in a chapter that contains twice as many words as it need do, make the interesting, and counterintuitive, suggestion that the later a function normally develops the less likely is compensation to occur after hemidecortication: in patients who underwent the operation in late childhood Kohn and Dennis found no difference in verbal intelligence between left and right hemisphere groups; and both groups performed almost perfectly on tests of visual closure and tactile form recognition; but those who had undergone right hemispherectomy were consistently inferior on map-reading and maze tests. Kohn and Dennis suppose that the skills involved in the latter tests are ones that do not normally mature until the second decade of life, whereas the skills for which recovery is apparently complete are those that mature earlier. However, the age at which a task can be successfully completed must surely be confounded with the, independently assessed, difficulty of the task: the tests of visual closure and tactile recognition may not have been delicate enough to reveal hemispheric differences and Kohn and Dennis may be among the surprisingly many who continue to be deceived by ceiling effects.

The lengthy catalogue of differences between the hemispheres in adults is extended in two chapters on olfaction and audition by Gordon. In commissurotomized patients he found that odours presented to the left nostril were better named (the olfactory nerve does not cross) than those presented to the right, whereas odours presented to the right nostril were more readily matched to the corresponding object. In another study, using the reversible hemispherectomy that is achieved by injection of amobarbitol into one carotid artery, he showed that singing was severely impaired after anaesthesia of the right hemisphere, tone being affected much more than rhythm.

Jerre Levy presses the now popular idea that the left hemisphere is analytic and sequential in its operation whereas the right hemisphere is synthetic and global, the two modes of function depending on incompatible neural organization; and Nebes offers some experimental evidence that the right hemisphere does excel in the combination of fragmented sensory information. However, Kinsbourne, in a concluding chapter, is quick to emphasize that the thought of the normal man is mixed in type and that it is the great cerebral commissure that allows this second-order synthesis.

Many of the early chapters of the book are pedestrian; many of the experimental results have already been published; and one chapter, that by Austin, Hayward and Rouhe, should have been deleted by the editors; but the latter part of the book, consisting of a series of chapters by Trevarthen and Kinsbourne, leaves the common rut of discussions of interhemispheric relations. Both these authors dismiss, for example, the notion of Gazzangia and Dimond that the brain of the commissurotomized patients consists of two independent channels. Trevarthen suggests that the function he has called ambient vision is not divided by commissurotomy: the commissurotomized patient can report verbally the gross position and movement of stimuli in the left visual field and can report the temporal order of a pair of brief flashes if one is delivered to each hemifield. In particular, Kinsbourne and Trevarthen both argue that the attentional resources of the two hemispheres are very far from independent and indeed propose that the corpus callosum is necessary to the normal distribution of attention. If the commissurotomized patient concentrates on a verbal task then he is unable to tap concurrently with his left index finger. Similarly, no evidence for parallel processing was provided by a visual search task in which one or two upper-case letters were briefly exposed in one half-field or one in each half-field: when a target letter was present on the left, the response was either absent or was of very much longer latency than responses made to targets appearing in the right field. If, instead of being allowed to respond with either hand, the patient was required to respond only with his right hand, then his neglect of the left side of the tachistoscopic field was absolute. Kinsbourne and Trevarthen attribute this extreme attentional bias to the loss of the equilibrating function of the corpus callosum. Commissurotomy cannot be understood simply in terms of disconnexion: 'Even more than an information transmitter, the corpus callosum is a regulator.'

In three concluding chapters Kinsbourne develops his theory of opponent neural systems linked by reciprocal inhibition. The principles elaborated by Sherrington for the spinal reflexes and by Hartline for sensory analysis are extended to the interaction between higher-order processes that may or may not be symmetrically represented. Reciprocal inhibition ensures that when incompatible decisions are simultaneously formulated, one will override the other and command the musculature. The most ancient of the reciprocal linkages, those mediated by the interhemispheric commissures, are traceable to primitive turning reflexes and reveal themselves in the contralateral neglect that follows unilateral damage in man and many other species. Damage to one polar element of a system will cause imbalances that generate symptoms that do not occur when both elements are approximately equally impaired. Both Kinsbourne and Bogen mention cases where a behavioural function seems to be destroyed by a focal lesion but is restored by a second lesion: after hemispherectomy in cats, forelimb placing on the opposite side is abolished, but the response can be restored by frontal lobectomy on the spared side; similarly, Sprague showed that the contralateral neglect produced by a unilateral, posterior lesion of the cortex could be abolished by excision of the superior colliculus on the side opposite the first lesion.

Kinsbourne and Smith's book should bring an end to incomplete accounts of commissurotomy in terms of simple disconnexion and should check incautious extrapolations to the normal brain. The disorders of attention that follow commissurotomy are emphasized, as are the extreme variability of individual cases and the abnormalities of cerebral organization that probably characterize these epileptic patients. But who then wrote on the dust-jacket: 'Callosal section makes it possible to study the unique contribution of each hemisphere in virtual isolation from the other'?

J. D. MOLLON

Frames and Cages: Repertory Grid Explorations of Personal Space. By ANTHONY RYLE. London: Sussex University Press. 1975. Pp. 148. £4.50.

This book is to be welcomed as the first on the subject of repertory grid technique since Bannister and Mair's *Evaluation of Personal Constructs* was published in 1968. *Frames and Cages* differs from its forerunner in many respects, but perhaps particularly in offering a practitioner's rather than an academic's approach to the study of grid technique.

Ryle starts by giving a superficial coverage of personal construct theory and then compares this with other theories, particularly object relations theory. This is a worth-while attempt which fails. It fails because of Ryle's incomplete grasp of the fundamental ideas of construct theory. In some cases this results in confusion (as in his attempt to explain the difference between 'construct' and 'element' in the organization corollary), in others, erroneous deductions are made which are then used to criticize the theory.

Next described is basic grid technique (mainly the rated grid), Slater's principal-components analysis and the interpretation of results. The following chapters on the interpretation of results in the light of psychoanalytic theory, the use of grid technique in the clinical situation and the case histories are of interest, the latter being particularly valuable to anyone new to the technique.

But this is not a book for anyone who wants to evaluate critically the method. Ryle makes no attempt to discuss the literature concerning types of construct, validity, reliability, different sorts of grid and analyses, or its many reported uses as a measure of change in psychotherapy. In many cases it is apparent that he is unaware that such literature exists. Also, in many of the examples, inadequate data (e.g. correlations, loadings) are presented to enable the critical reader to follow how he reaches the conclusions he does. This is therefore a useful book for the psychologically unsophisticated reader who is looking for a general account of what grid technique is all about, but is unlikely to meet the needs of most psychologists.

FAY FRANSELLA

Clinical Psychology: The Study of Personality and Behaviour. By SOL L. GARFIELD. London: Arnold. 1975. Pp. 461. £8.50.

The book is parochial enough to merit re-titling *American Clinical Psychology* but nevertheless much of its content is relevant hereabouts. It is a workmanlike attempt to cover the spectrum of clinical concerns ranging from the kinds of personality theory and research areas which have influenced the field, through assessment methods (test and interview) to therapies (psychotherapy, behaviour therapy, group therapy) and community work. Each topic is at the outset viewed historically which gives the book a broader perspective than is usual. This perspective is further enlarged by specific discussions of the context in which clinical psychologists work – their professional and intellectual warfare with medicine, their training problems, their ethical dilemmas and so forth. The language of the book is clear, if unstylish, and the survey of literature adequate if too restricted to self-labelled 'clinical' sources.

The author's respectful striving for the middle ground of clinical psychology has curiously thrown him off balance at times. Like many psychologists writing about a rapidly changing field he has opted for the chronological as well as the political centre. Thus his weighting of topics is often mid-sixties rather than mid-seventies. For example, his lengthy treatment of projective tests such as the Rorschach and trait-diagnostic tests such as the MMPI is not reflected in the assessment habits of most contemporary clinical psychologists. On the other hand, being American, being the Garfield of Bergin and Garfield (*Handbook of Psychotherapy and*

Behaviour Change) and being determined to include a mention of all ongoing ventures he has set his volume at about the level at which clinical psychologists are working in the field (i.e. he includes material on psychotherapy, community work, group strategies, experiential teaching methods) and about ten years ahead of the level being taught in British university training courses.

Garfield deals at length with the way in which clinical psychologists have changed their role. He details the shift from the mental tester of the late fifties (a time when the profession came close to boring itself into extinction) to the therapeutically oriented contemporary psychologist who is additionally moving towards a 'social psychologist to the institution' – 'preventative psychology for the community' role. It is encouraging to note that when this saga is set out in detail it reveals not only the confusions of psychologists but their capacity to develop new themes out of their confusions and to respond to the increasingly varied demands made on them.

The book can be strongly recommended as an undergraduate text, particularly because it intelligently relates the problems of abnormal psychology as a discipline to the problems of clinical psychology as a profession.

D. BANNISTER

Understanding Conflict and War. Volume 1. *The Dynamic Psychological Field.* By RUDOLPH J. RUMMEL. New York: Wiley, Halsted Press Division. 1975. Pp. 342. £10.35.

A reviewer of this book at this time is in the tantalizingly impossible position of an opera critic at a world première having to write his piece after hearing only the overture. Over the next few years more volumes (four or so) are promised. And the series needs to be viewed against the background of *Applied Factor Analysis* (1970) and *Dimensions of Nations* (1972), in the second of which Professor Rummell, a professional political scientist, described his application to international relations of the technique he expounded in the first. In the present volume his aim is no less daunting than to analyse man's behaviour in society so as to lay the psychological and philosophical foundations for a unified theory of war which would help to bring about its elimination.

The keystone of his approach is field theory, which is surveyed in Part I. Although rejecting the physical sciences as a model for the explanation of man, he borrows the field concept and reformulates the mathematical functions to portray social phenomena. His analysis seeks to remedy the weaknesses in existing field theories of man. Next comes an extended account of perception as a dynamic confrontation between internal and external vectors. He offers here a description of perception in terms of latent functions and a philosophy of reality derived from Ushenko. Part III concerns expectations and behaviour. To describe the space of human behaviour he proposes a behavioural equation modelled on multivariate psychology – in particular Raymond Cattell's 'specification equation' giving the dependence of behaviour on personality in a situation. Then comes motivation, his hybrid approach being 'a mixture of the teleological model and the existential view, with the tension-release model playing a subordinate role.' The empirical basis for the detailed picture is Cattell's research, and other motives which do not immediately fit into this picture (like aggression, curiously enough) are reinterpreted ('rotated' might be the *mot juste*) until they do. Part V broadens the perspective to show man's dynamic field in its wider sociocultural and ecological setting, while the last two sections are predominantly philosophical. In Part VI, Rummell attacks 'the ultimate question', the free will–determinism controversy, because he believes 'To create a just world, we must *affirm* freedom'. He adopts a revised version of Kant's solution because it lets us have our cake and eat it too, harmonizing humanism with science. In the last part he sketches his philosophy of 'intentional humanism', a middle-of-the-road, man-centred world-view emphasizing the complementarity of science and humanist intuition.

The whole thing is the stupendously ambitious venture of a polymath, and reservations about so grand an enterprise are inevitable. To dispose first of the lesser ones, there is the tiresome blemish that minor inaccuracies are fairly prevalent in the text, due perhaps to poor proofreading. It is unfortunate, too, that 'cognitive dissonance' is used as a blanket term covering the inconsistency theories, and Rummell's attempt to translate them into his own terminology is more confusing than illuminating. And he has a passion for spatial representations, such as diagrams of psychological space, which are uncannily reminiscent of Lewin's life-space diagrams which he has earlier criticized. His Kantian treatment of free will, whether intrinsically convincing or not, smacks of wishful philosophizing when one reads that it was adopted because 'Within the context of my field approach to reality, my perspective realism and pragmatism, Kant's solution is the most congenial.' Nor is it open to the psychologist reader to pass this buck with relief to the philosophers, for Rummell's equation (unlike Cattell's) contains a term representing will. We may feel this is multiplying hypotheses beyond necessity, but in any case what empirical nature is to be ascribed to such a term, how is it to be quantified and measured? A final quantitative question: what is the optimum length nowadays for a *magnum opus*? This first volume has only a tiny, incidental amount to say about conflict and war, but plenty of coffee-house psychology, philosophy, and sociology. Its monumental spadework is in places uneconomical, repetitious, and over-elaborate, so that one ends wondering whether the enterprise might not be successfully accomplished in fewer than four further volumes.

One is also left with more favourable impressions. A richly stocked background arises from Rummell's anthology of other men's works: he provides well-conducted tours of thought marked by clear logical development, penetration, and occasional charm and humour. His integration of the work of Lewin, Cattell and Sorokin looks a valuable contribution to the development of field theory in the social sciences. And whether or not this is a time for polymathy, his encyclopaedic, man-centred perspective performs the service of reviving the big, perennial questions in the contemporary context. I suspect that many, having overcome some initial trepidation, will find the book an experience they have largely enjoyed even if they are not wholly persuaded by the argument.

ALFRED FLOOK

The Social Psychology of Religion. By MICHAEL ARGYLE AND BENJAMIN BEIT-HALLAHMI. London: Routledge & Kegan Paul. 1975. Pp. x+246. £5.95.

This is the second edition of a book first published in 1958, but so much more research has been done since then, which is reviewed here, that it is virtually a new book. Various chapters review the relationships between religious behaviour and factors such as age, sex, personality, social and political attitudes and other economic factors. Other sections consider the obverse influence of religious behaviour on factors such as personality adjustment, sexual and other social mores, though the authors frequently point out that it is often difficult, if not impossible, to discern the causal relationships between religion and these other factors investigated. Finally, there is a more theoretical last chapter discussing some of the many theories that have been put forward to explain, firstly, the origins of religious behaviour; secondly, their continuation in the form of belief systems; and thirdly, the effects of religious behaviour on all other aspects of life.

Most of the material studied in the early chapters is statistical and correlational, and therefore inevitably, its cool and critical review by the authors has a slightly remote and dry quality. Although the last chapter does have some statistical data, much of it is on a different level from the rest of the book. It reviews some of the famous psychosocial theories of religion such as those of Freud, Marx, Durkheim, etc. Most of these theories do have a certain amount of empirical evidence to support them, so that, not surprisingly, it seems that a large number of psychological and social factors influence the numerous and various beliefs and practices subsumed under the umbrella of religion.

The authors comment that despite the apparent decline of religion in the West as a system of beliefs, and the fact that it seems to have little psychological effect for good or evil on individual behaviour and the sense of personal wholeness that is usually equated with sound mental health, nevertheless its social effects and educational influence are still powerful. Religion as anthropologists see it in primitive societies is usually an even more powerful and widespread psychosocial force. Though the authors take an open and objective view of religious behaviour, they also point out the limitations of psychology, even though it has come a long way since the days of early behaviourism, with research now being carried out into conscious experiences, creativity, etc. Nevertheless, many psychologists (including the authors) feel that the subject needs to extend its philosophical base still further and, provided that proper research methods to go with such an extension can be elaborated, the relationships between psychology and religion may be very different and less one way. Anthropologists, such as Nadel, point out that religion may influence cosmology, moral values and social cohesion, and be closely concerned with the production of certain specific mental states, usually of an ecstatic or mystical character. As can be seen in Confucianism, such a widespread influence on society and the individual need not be centred on a personal God, but thousands of years of the Greek–Jewish–Christian Western tradition make it hard for us to envisage a religion that does not place individual personality at the centre. For us the aphorism of the contemporary existentialist theologian, Bultmann, holds good – 'The question of God and the question of myself are identical.'

DESMOND POND

Community Mental Health Programmes in England: An American View. by N. ROSENZWEIG. Detroit: Wayne State University Press. 1975. Pp. 282. $13.95.

The book stems from two visits to England made by Dr Rosenzweig in 1969 and 1971. It is aimed primarily at the American market, with descriptions of a variety of community mental health service schemes in England which can be compared with the mental health centre services provided in America. It is an interesting, well-written and comprehensive account of the ways in which these services can be delivered, emphasizing the importance of individuals to the success of such schemes and the problems that are encountered.

The first third of the book contains concise accounts of some of the more general issues involved, such as the basic differences between the services in each country, the National Health Service and the then impending reorganization; the psychiatrist and other professionals concerned with treatment, and the way in which services are delivered. The last chapter in this section concentrates on recently developed arrangements that are part of the attempt to reduce the number of psychiatric beds in institutions. It provides a good introduction to the second part of the book which describes some of these efforts in much more detail.

The way in which Dr Rosenzweig describes these services is very appealing. He takes us on a conducted tour of the schemes he visited, operated by psychiatric hospitals, local authorities, those run jointly by psychiatric hospitals and local authorities, and those provided in general hospitals and by voluntary organizations. In addition, he describes the unique services provided by the Tavistock Centre and the Institute of Psychiatry and Maudsley Hospital in London, and the Institute of Family Psychiatry in Ipswich. The final chapter is devoted to an in-depth study of the services provided in the London Borough of Camden.

The author makes it clear that the places he visited should not be taken as representative; and indeed most of his observations stem from schemes that have been relatively successful. However, he touches on a variety of points that are obviously not peculiar to these schemes, such as problems of coordination between and within the different types of organizations involved, the importance of voluntary organizations and the implications of the Seebohm reorganization in social work for community mental health services. However, insufficient attention is given to the way in which resources are obtained and, inevitably, certain terms such as Medical Officer of Health and Local Health Authority are already out of date.

It is unfortunate that the book was written just prior to the reorganizations of local government (outside London) and the health service. In his concluding comments Dr Rosenzweig makes a very good attempt at speculating about the future of community mental health services, and warns against, for example, too rapid a transfer to a community-based service, which happened in America when the community services were insufficiently developed to cope with the demands subsequently made on them. He points to the need for more general medical practitioner involvement via the primary care team idea which links the GP with the social worker, community nurses and the psychiatrist, and for more effective collaboration between the health service and local government. His comments about the need for psychiatry to become better established within the general hospital specialities are absolutely right.

It would have helped to round off the book if Dr Rosenzweig could have talked in a little more detail about the new health service and arrangements for collaboration with local government services, but unfortunately this was not possible at the time the book was written. However, he has written an interesting account of community mental health schemes that can, and should, be read by individuals from the various disciplines involved.

S. GRIFFIN

Varieties of Residential Experience. Edited by Jack Tizard, Ian Sinclair and R. V. G. Clarke. London and Boston: Routledge & Kegan Paul. 1975. Pp. viii +290. £6.50.

At any given time, more than one in fifty of all children and young people in England and Wales are in some form of residential accommodation, because the problems they present (ranging from need for overnight accommodation because mother has been taken ill, to homicide) are such that they cannot be dealt with in the community. They represent the most damaged, deprived and disruptive of our youngsters. Vast sums of money and numbers of people are devoted to their care, custody and treatment, with results which are not always clear or encouraging. We need to have a good deal more coherent information about what is happening to them.

This book is welcome if for no other reason than three well-known British researchers (two psychologists and a sociologist) have recognized the need and brought together a number of research contributions which seek to highlight problems and issues in 'residential experience' – as it turns out, more the researchers' experience of evaluating residential establishments than the residents' experience of their settings.

The editors disarmingly admit that the eleven chapters in their book present an uneven picture of residential provision, because of the patchiness of the research in this area. Their own introductory chapter rejects the 'steampress' or totalitarian model of institutions exemplified by Goffman in favour of a comparative approach and an emphasis on institutional diversity. The only attempt at producing a conceptual framework is made, in two chapters, by Millham *et al.* Their presentation is complex and sophisticated, and although it takes no account of modern psychological approaches to socialization by, for example, Bandura, Mischel or Sarason, psychologists would be the poorer for not reading it.

The remaining chapters present results of empirical research, of a largely descriptive and survey type, into nurseries, units for retarded and autistic children, approved schools and probation hostels. Peter Moss provides useful information and statistics on the structure of residential care; Jack Tizard, Barbara Tizard, Lawrence Bartak and Michael Rutter discuss their respective researches on homes for the retarded, nurseries and units for autistic children, each in the best tradition of comparative studies; Kevin Heal and Pat Cawson take a particularly simplistic look at organization and change in former approved schools; Tadeusz Grygier's paper on the development of a predictive instrument for use with youngsters in correctional institutions in Canada, seems out of place in that it is neither the result of work in Britain, nor is it comparative in the same sense as the others; Ian Sinclair's work on probation hostels, highlighting the importance of environmental and interpersonal factors in directing behaviour receives another impressive presentation; Ron Clarke and David Martin continue his theme in emphasizing the role of environmental variables in affecting absconding amongst approved-school boys. Their chapter is exemplary in presenting complex data in an uncomplicated fashion and in squarely facing the implications of their research.

The book raises issues of wider psychological interest than indicated by the title. It also highlights the strengths and weaknesses of this form of comparative research. The chapters are valuable in showing the quality of the best work that is carried out on institutions and the enormous complexity of the issues involved. But the book as a whole does not hang together, neither in approach nor in subject matter. Given the price of the book and its distinguished editorship, it would have been reasonable to expect at least an attempt at a synthesis or an empirical overview based on these and other researches on institutions. Such an attempt needs to be made if research is to be of use in shaping our future residential provisions. In the meantime, one should be grateful for this book in providing some of the best material on which a synthesis would be based.

M. S. HOGHUGHI

The Gifted Child. By Cyril Burt. London: Hodder & Stoughton. 1975. Pp. x+214. £2.60.

A book on gifted children by Cyril Burt, published four years after his death, must be to many an unexpected addition to psycholcgical literature. Of course, Burt was interested in the gifted from the outset of his career, since one of his assignments when he commenced his work with the LCC in May 1913 was to help in the ascertainment of bright children for scholarships. So he began collecting material on the gifted very early in his career, writing up his conclusions in several reports and articles, in particular a long contribution to *The Yearbook of Education* for 1962. Assembling this earlier material in book form was in fact the last major job of work which Burt, who remained productive and professionally active up to the last weeks of his long life, managed to complete. As Dr Charlotte Banks notes in her brief introduction to the present work, before Burt's death the typescript in its final form was complete apart from a few tables. So, though it has taken four years to achieve publication, the book may be regarded as having received his imprimatur.

A book on the gifted child by Cyril Burt has potentially a twofold value, as a contribution to the psychology of the gifted child, and for the light that it throws on its author's psychological views. It is, I think, on the latter count that it is likely to prove of principal interest. As a contribution to the study of the gifted it is of somewhat limited value; firstly because it does not in fact contain much new material. Burt's findings have been published before; there is nothing essentially new in the present volume, and a good deal of it really consists of padding. Secondly, the book takes little notice of the considerable volume of recent work on the gifted. Terman's classic study is often referred to, and also Parkyn's New Zealand investigation, but the researches and the literature of the 1960s are for the most part ignored. So the book is hardly an up-to-date guide to the psychology of gifted children.

It does, however, contain a good deal of material relevant to an assessment of Burt's contribution to psychology. The first two-thirds of the book are a sort of prolegomenon to the problem of the gifted, in which Burt summarizes his views, already reported in numerous books and articles, on the nature and assessment of intelligence, the evidence for the concept of intelligence, the general principles of genetics, and the evolution of the human species. It is useful to have this material set out succinctly in one place, though it is not likely to convince those who are not already satisfied by the evidence. The difficulty level of the different chapters is somewhat variable, and in particular the chapter on multifactorial inheritance may prove unduly mathematical for some of the ordinary teachers for whom the book is primarily intended. Since Burt died, his genetical findings have become a source of considerable controversy. Kamin, in his recent book *The Science and Politics of IQ* (1974), has damagingly exposed inconsistencies and improbabilities in Burt's published kinship correlations. Burt's posthumous book unfortunately does nothing to resolve the difficulties. The tables have been reprinted in the form in which they have previously appeared, and the discrepancies remain as before. The problems are more deep-seated than Dr Charlotte Banks suggests in her introduction, and, as the raw data have disappeared, they can never be resolved.

In some ways the most valuable chapter in the book is the last chapter on the education of the gifted. It is brief, but Burt here brings to bear his considerable powers of judgement and his practical wisdom. For Burt was not merely one of our most eminent academic psychologists, he was also, and perhaps first and foremost, an applied psychologist, and his feet were firmly planted on the ground of practical realities. What he has to say about the education of gifted children is sensible and imaginative, though it needs supplementing in the light of recent practical experiments.

Though some unpublished manuscripts still remain among Burt's papers this is likely to be the last book we shall read from a distinguished psychologist who wrote prolifically, but always with clarity and erudition.

L. S. HEARNSHAW

ognition and the Symbolic Processes. Edited by WALTER B. WEIMER and DAVID S. PALERMO. New Jersey: Lawrence Erlbaum. 1974. Pp. ix + 450. £10.00.

First things first: the book contains the proceedings of the Conference on Cognition and the Symbolic Processes, held at Pennsylvania State University in October 1972. In addition, there are three papers included in the volume which were written after the conference, partly as a response to it. There are papers on cognition versus conditioning (Brewer and Dulany), on Gibson (J. J.) and why he was right all the time (Mace and Turvey), on comprehension and understanding (Bransford and McCarrell and Franks), on where psycholinguists should go from here (Ross and McCawley), on I know not what (Shaw and McIntyre and Halwes), on wholism and atomism in neuropsychology (Anderson), on vocabulary growth (Miller) and Weimer as general over-seer. Structually the book seems to consist of papers by a dominant and consensually validating core (mostly Brewer, Dulany, Shaw, Halwes, Anderson and Weimer) with papers by out-group members (Ross, Miller, Bransford) placed on the periphery. This means a lively, sometimes rude, and usually fruitful interchange in the discussions of periphery papers, but a lot of unnecessary social grooming for the core ones (except for the splendid intervention on the part of an unidentified speaker in the audience for Brewer's paper which forced Brewer to execute a scarcely dignified retreat from his rather simple-minded thesis that There Is No Convincing Evidence For Operant Or Classical Conditioning In Adult Humans). The discussion of the Shaw/McIntyre work is particularly poor here since I have a more than sneaking suspicion that someone less in awe than the discussants might have detected more than a whiff of the obvious, the trivial and other peoples' ideas behind the opaque surface of the paper (Weimer in his later discussion seems to say that the lack of clarity in Shaw's contribution is *a priori* a virtue – well, mm, perhaps).

Also, there is too much of Kuhn in the book, many of the writers having convinced themselves by that impeccable brand of logic that persuades little boys that if they smoke cigarettes then they will automatically become men since only men are allowed to smoke, that mere disagreement with their elders (and sometimes betters) implies that they have discovered the New Paradigm, that normal science, as we have all come to know and love it, is overthrown and that liberating revolution is just around the corner. Happily, other people strike a sourer or at least a more cautious note. Finally, and seemingly also part of the Kuhnian movement, is the use of hindsight as a substitute for rational criticism in one or two of the papers, particularly in the rather gratuitously rude attack on Miller by Weimer at the end of the book. OK, perhaps we can concede Weimer's point that Miller has not always been right in the past (however 'right' is defined here), but that is no reason for assuming that what Miller suggests now is necessarily incorrect as well. Perhaps other and less contentious criteria are needed for judging the value of present work, such as how insightful is any synthesis offered, how well does it accord with evidence in other areas, etc., rather than on the basis of some spurious kind of academic track record.

Articles which I found interesting and useful were those by Miller (elegant and lucid as ever), Bransford *et al.* and the ones by Turvey and Mace, although the former seemed to have more than a touch of the lecture-room about it. Other readers will, I am sure, thank Brewer for his exhaustive *Psychological Bulletin*-type review of human conditioning studies, while the resulting interchanges between him and Dulaney point up a great many of the problems and, some would argue, the unrealities of the great cognition/behaviourism debate. Anderson's rather extravagantly written note on neuropsychology should stimulate a parallel debate in psychology, one which is long overdue. One or two sceptical notes at the end, viz. many of the papers give me an overwhelming feeling of *déjà vu*: Skinner versus the rest, Gibson ditto, the Würtzburg school for ever, etc., is it really evolution or is it merely revolution? What exactly is a cognitive approach to anything? How can one operationalize all of those high-falutin' notions and get a crucial experiment from them, or even any kind of experiment at all?

A. D. LOVIE

The Tell-Tale Eye. By ECKHARD H. HESS. New York: Van Nostrand Reinhold. 1975. Pp. 250. $10.95.

This is a glossy, chatty, semi-popular book about the work by the author and his associates on pupil dilation and constriction. Research by others in the same area is mentioned briefly. Hess, as is now well known, observed in 1960 that male pupils dilate to photographs of nude females, while female pupils dilate to nude males and to babies. This led to a number of studies by Hess and his group (Polt, Seltzer and Schlien), for example, on the pupil dilation of homosexuals, the dilation of pupils while doing difficult mental arithmetic (confirmed by Kahneman), dilation to diagrams of eyes and dilated pupils (confirmed by Coss), and dilation to emotive words. It was found that pupillary responses to photographs of Johnson and Goldwater changed after reading political propaganda. Hess also reports his adventures in the advertising world, where he found that pupillary responses to still advertisements predicted sales better than did verbal responses, and he was able to find the points of greatest interest in TV commercials – though he was not able to test whether these commercials were effective.

A number of the findings have not been replicated. Most research workers have failed to find pupil constriction in response to unpleasant stimuli. Their work is dismissed but not discussed in detail. There are conflicting findings on the correlation between pupil dilation and verbal reports of interest or arousal. Some studies have found that dilation is correlated with introversion, others with neuroticism but not with introversion. Work by the Hess group on the relation between pupil dilation and other measures of arousal obtained rather complex relationships whose interpretation is obscure. Materials are provided for the reader to repeat an experiment in which it was found that subjects drew larger pupils for smiling faces than for frowning faces, but it is not clear what we learn from this. A number of these studies have not been published in journals, used very small numbers of subjects (4–6) or were not designed in a way that would enable important hypotheses to be tested.

There is a chapter on the history of pupil research, which does not mention work in ethology, and an appendix on techniques of measuring pupil response.

This book does not provide scholarly account of work in this field: Goldwater's paper (1972) is very much better in this respect. There is also a complete absence of any theoretical or explanatory material. This could be at the evolutionary level, as provided by the recent work of ethologists like R. G. Coss and M. Scaife, or it could be at the physiological level. All we have is a series of rather simple experiments, and the assertion that arousing stimuli produce dilation, while unpleasant ones produce constriction. I think that Hess could produce a much more scholarly and interesting book than this one, and I hope that he will do so. MICHAEL ARGYLE

REFERENCE

GOLDWATER, B. G. (1972). Psychological significance of pupillary movements. *Psychol. Bull.* 77, 340–355.

The Perception of Color. By RALPH M. EVANS. New York: Wiley. 1974. Pp. ix + 248. £8.10.

As every schoolboy knows, in the case of human colour vision the magical number is three. All colours can be matched by mixtures of three others, either as lights or pigments [the fact of trichromacy from which Thomas Young deduced (correctly) that there are three colour-sensitive mechanisms in the retina]; for such matches, three psychophysical variables (dominant wavelength, purity and luminance) are necessary and sufficient; and the three perceptual variables of hue, saturation and brightness define completely the qualitative and quantitative aspects of colours as they are seen. It is the last of these widely held beliefs which the author of this book is concerned to challenge. While fully accepting trivariance at the receptor level, he argues that there are five perceptual variables of colour, reducible to three only under special circumstances; the additional variables are 'lightness' and 'brilliance'.

Now the assumption of three perceptual variables rests on the opinion of no less an authority than Helmholtz. Having established from his own researches and those of Clerk Maxwell and Grassmann that all colour stimuli could be grouped into families (now called 'metamers') which look alike but differ in spectral composition, and that each family could be specified by three independent psychophysical variables, Helmholtz proceeded to the 'logical' assumption that the colours seen could also be specified by three perceptual variables. However, as Evans points out, this assumption is reasonable only if one also assumes that the appearance of the stimulus is controlled entirely by the stimulus itself. But this is true only in the special case of the 'isolated' stimulus, i.e. when a single, uniform, stimulus is the only one sending light to the eye: under these conditions (considered in Section I of the book) there is a fairly well-established relationship between the three psychophysical and the three perceptual variables. Once it is recognized that the appearance of the stimulus itself is changed by the presence of neighbouring stimuli (Section II: 'Related Stimuli'), it becomes unwarranted to assume that these changes must occur in the *same* perceptual variables that are controlled by an isolated stimulus. The author suggests that 'Only a persistent desire to keep the system three-dimensional (so that it can be visualized?) can explain the circumlocutions that have been resorted to, to make it so appear'. 'Lightness' and 'brilliance' are contextual variables, produced by the presence of a second stimulus, and are *in addition to*, not modifications of, the three perceptual variables produced by an isolated stimulus. In order to account for the five perceptual variables, Evans finds it necessary to add to the usual luminous efficiency and colour-mixture functions of colorimetry a further function which he calls 'chromatic strength', which depends on the relation of the psychophysical variables of the stimulus to those of the surround.

The case for these additional variables is one that psychologists and artists, if not colorimetrists, are likely to find plausible. Unfortunately, it rests at present on a rather flimsy experimental basis, mainly on the results of exploratory studies using only two observers. Moreover, although the discussion extends beyond the confines of conventional 'stimulus colorimetry', to include what Evans calls 'perceptual colorimetry', it is still restricted to the hypothetical 'Standard Observer' of the international C.I.E. system of colour specification. Many psychologists will remain sceptical as to whether even five 'basic' variables, supplemented as necessary by the additional cognitive concepts (e.g. 'frames of reference') to which the author appeals in Section III ('The Observer and the General Stimulus') will suffice to account for the colour perceptions of real observers. And some extremists among them may feel that the author's classical and essentially Helmholtzian approach to problems of colour perception (hinted at in the Introduction and reflected in the general plan of the book, though never made explicit and defended) is misguided and should be discarded in favour of the more radical alternative of attempting to develop a 'perceptual psychophysics' of colour along Gibsonian lines.

Despite the limitations of his experimental data and his reluctance to engage in theoretical discussion, the author has nevertheless succeeded in producing an original and challenging book with far-reaching implications for research on colour specification and perception. In particular, it suggests the need for re-examining and extending earlier work on 'stimulus colorimetry' and three-variable 'colour-order' systems, such as the Munsell. More generally, it indicates that far from being the closed subject it is often assumed to be, colour perception is wide open for further research, especially at the level of pattern and object perception. If students of perception can manage to overcome their initial disappointment on discovering that *The Perception of Color* contains no illustrations in colour, and are not deterred by the (visually) rather uninviting text, they will find this book demanding but very rewarding to read.

G. W. GRANGER

Architectural Psychology. Proceedings of the Lund Conference. Edited by R. KULLER. Stroudsburg: Dowden, Hutchinson & Ross. 1973. Pp. 450. £13.00.

Along with other such clichés, the estate agents described our first home, an Edwardian attic flat, as 'desirable'; our second, a Victorian terrace house as 'functional' and our third, an Elizabethan semi-detached villa as 'convenient'! My wife described them respectively as 'depressing', 'frustrating' and 'cramped'. Upon reflexion, for me the flat was 'about the best I could afford at the time', the house was 'located close to the University where I worked' and the villa had 'a pleasant garden for the kids to play in'. Whilst reading this book I found myself wondering whether those responsible for planning, designing and building these three dwellings might have reconciled our different opinions had they found themselves amongst the 83 delegates to the Second Architectural Psychology conference held at the University of Lund, Sweden in June 1973, the proceedings of which it reports. Aesthetic, economic, geographical, social and variously behavioural – surely reconciliation of our various criteria would seem far beyond the resources of a single study group. Nevertheless, the comfort, well-being and general quality of these three phases of our lives would undoubtedly have been enhanced by application of the principles developed in this volume. Referring to *New tendencies in lighting research* and pointing out that: 'From the soft lights of the intimate night club to the harsh glare of the interrogation situation, there is a rich variety of ways of manipulating the flow of light to influence interpersonal communication' the lighting engineers would have demonstrated how our flat could have been improved from development of their: 'greater insight into the types of impression that can be fostered or reinforced by various lighting systems in a behaviour setting'. If this proved the case then their conclusion suggesting that: 'Perhaps this is a step towards illuminating engineering becoming recognized as an aspect of social psychology' seems fully justified. Attempting to delineate *Problems and theoretical models in architectural psychology* it was agreed that: 'Contemporary repetitive mass-produced apartment blocks have contributed much to the present welfare both directly by providing better accommodations for more people and indirectly by increasing the efficiency of the working force'. Echoing our own criteria, however, it was also acknowledged that: 'these facts cannot be used to suppress the public critique of modern buildings for being dull and depressingly monotonous'. With reference to our own terrace house it can scarcely be denied that: 'understanding of the perceptual requirements to be met by the physical environment would seem fundamental' and that: 'experiments measuring aesthetic factors may provide a more rational basis for evaluation and design of industrialized housing', than the criteria appealed to traditionally by architects, builders and planners. Discussion of *Applied research in the field of architectural psychology* revealed: 'a central concern with improving or making better the quality of the built environment. Yet, there has been relatively little attempt to clarify what it is which makes one environment better than another nor even to define precisely what may be meant by the term "better". Underlying this is a concern with the evaluation of buildings'. Hopefully, such evaluation, 'will enable architects, their associates and clients to determine whether existing or proposed buildings are likely to be better or worse than any available alternatives'. Moreover, intending shortly to move from our semi-detached villa, considerations referring to the *Place of theory of architectural psychology in tomorrow's planning* might help us in choosing a new house. Having complained at length: 'Over the lack of communication between architects and psychologists, i.e. the discrepancy between knowledge about man's reactions to his environment on one side, and knowledge about those physical and social realities which form the environment on the other', the objective remains: 'to chart the relations between what we perceive and how it is perceived. There is a need for useful models to relate stimuli to perceptions. Such models are a necessary prerequisite for the much demanded cooperation between consumer, psychologist and architect in the process of planning'. Since delegates to the Lund Conference heralded the promise of such cooperation, we may yet find a desirable, functional and convenient place to live.

GERALD H. FISHER

Soviet Psychology: History, Theory, Content. By J. McLEISH. London: Methuen. 1975. Pp. xii + 308. £6.95.

A better title for this book would have been *Soviet Psychology: The Party Line.* For Professor McLeish has swallowed this line completely, and hook and sinker too. And he follows it through all its twists and turns, whether the angler is Marx himself, or Lenin or even Stalin. He faithfully records all the good reasons why the Party successively decided that this or that tendency in psychology was an 'idealist' (or 'vulgar mechanist' or 'bourgeois' or 'eclectic') deviation from the correct path of dialectical materialism (a purgative process which was carried through so thoroughly that for about 15 years, from 1936 to 1950, there was hardly any psychology at all in the Soviet Union!). It is difficult to tell whether he believes that *all* of these decisions were correct, since he writes throughout in a curious kind of *oratio obliqua* which never allows one to determine whether he approves of what the Party said and did or is merely telling you about it. But if he has any criticisms of the way science is kept subservient to Marxist–Leninism in the Soviet Union he keeps them pretty much to himself. He appears, for example, to have no doubt that to decide the truth or falsity of a particular psychological theory by whether or not it conforms to the revealed truth contained in the dialectical materialist canon is a perfectly reasonable way of conducting scientific affairs. There is much talk of 'practice' (a notable Marxist 'criterion of truth') but none of empirical, much less experimental, data deciding the issue between rival theories.

Professor McLeish does not even subject the claims that this or that approach was or was not in conformity with Marxist–Leninist dogma to much critical scrutiny, though this is the touchstone used to decide everything. For example, he accepts without question that Pavlov's theory of conditioned reflexes, which was canonized by the Soviet Academy of Sciences in 1950, 'provides the materialist context for the psychological study of the organism' and is in perfect accord with the fundamental principles of Marxist–Leninism. That, certainly, is the official Soviet view of the matter. But one would have expected him at least to have considered whether the Party needed Pavlov because Pavlov was a good Marxist or because he was a good scientist. Gustav Wetter, in his detailed book on 'Dialectical Materialism', concludes that Pavlov could just as easily have been regarded by the Party pundits as a 'vulgar mechanist', had they wished to hunt him out as a heresiarch. As it was, they reserved this epithet for some of his less brilliant followers.

In short, if you wish to find out what, actually, Soviet psychology has achieved or has to offer to the rest of us, this is not the book for you. (It would be very difficult to describe *any* empirical discovery about behaviour from the information contained in it.) But, if you are interested in a well-written case study of how the official philosophy of one of the world's great theocracies makes its science go through the hoops, it makes fascinating reading. And, having read it, you will be better placed to understand some of the more esoteric (to the Western reader) passages in Soviet psychological papers (the ones that quote from the works of J. Stalin, for example). In addition, there is an interesting account of the pre-Soviet social and philosophical climate in Russia and the way in which this continues to influence the contemporary scene. And, not least, you will learn quite a lot about Marxist–Leninism (which McLeish presents sensitively and lucidly, as well as with great sympathy), and its one saving grace: that, if it had only avoided fossilization as a state religion, it would have been a wonderfully well-balanced starting point from which to set out on an investigation of Man and his place in the natural world (compared, say, to the beliefs they hold in Rome or California). For it is indeed true that all those proscriptions of the wilder deviations (from Freud to the 'radical' behaviourists) have left Soviet psychology much saner than most of what passes under the same name in the West. One just wonders how far such sanity can take you along the paths of scientific discovery if it is believed that 'empirical studies arise out of this context of agreed assumptions and *cannot, in principle, contradict them*' (my italics: Professor McLeish doesn't think the remark worthy of them).

JEFFREY A. GRAY

Measurement of Human Resources. Edited by W. T. SINGLETON AND P. SPURGEON. London: Taylor & Francis. 1975. Pp. 368. £8.50.

The authors state that this book is about how total human resources may be moulded to increase the community's capacity to increase its wealth via the intelligent efforts of skilled responsible individuals. This paraphrases part of the introduction but illustrates the target they have set themselves. It is not in the nature of a book of contributions from various authors that such an ambitious target could be reached, but does it make progress towards it?

The contributions arose from a NATO symposium in Lisbon in 1973 and in the introduction the editors discuss the difficulties for the various disciplines to cooperate. With differing objectives, measuring methods and criteria it is not easy to build upon the results from one area of study in another. The book records the presentations from workers in psychology, manpower studies, education, selection, placement, career guidance, systems design and ergonomics. Unfortunately, although their differences are revealed we must wait for Burke's paper at the end of the book before some attempt at synthesis is made.

The first section of the book concentrates on the problems of structure and classification with three papers by eminent workers who have considered taxonomics: R. M. Gagné, R. M. Miller and E. A. Fleishman. The first two authors pursue problems of training and the role which ordered classifications of the multiplicity of variables associated with learning can play to aid the development and understanding of training and its results. Fleishman's chapter views research into human performance more widely, being concerned with attempts to find general factors, the influence of which can be recognized in a wide range of performances. He presents a very persuasive case for the work at AIR, but in common with his co-authors there is scant reference to work outside the US. It may be that the importance of this area is little recognized elsewhere but the discussions of other contributors do not bear this out.

The second section is the largest and is concerned with the development of resources. Some of these review the field whilst others discuss a particular aspect or technique. Each paper is of interest in itself but they do not cohere to present a structure for human resource development. The early papers deal with different aspects of the educational process. An interesting review by Butcher of resource use in British education is followed by two papers on specific studies in education. We then move on to career guidance, where three papers consider three bits of the subject, but no paper continues the argument from education to career with a competent review. In logical sequence, we then arrive at training, where Duncan presents his now widely known developments in task analysis and training. In spite of the able presentation this reviewer would have wished for more continuity via a discussion of training which saw it as wider than operator skills.

Two papers on process control skills include one by Bainbridge which is amongst the few which provide new material. She presents a technique for analysing the performance of people engaged in decision-making tasks from commentary provided by them whilst doing the job. The results are a step forward in the reporting of working behaviour so that the use by the worker of routines and memorized information can be studied and analysed.

The next four papers in this section are scattered across the resource field. One reviews some problems in the analysis of work. The second appears mistitled since it reports an interesting study of performance using a wide range of methods rather than providing a review. The third describes a study of perceptions of job stress and the fourth considers results from a study of employee involvement. Each is of interest in itself but the link is tenuous. The link is broken by the next papers, where an exploratory study of some problems in unemployment is followed by a paper on strategies for leadership utilization. This is wider in its discussion, presenting an area of study of undoubted importance for resource utilization, but detached in its impact on the book as a whole.

The penultimate section concerns human resource planning. Here are two descriptive papers concerning manpower planning and allocation in the British Civil Service and in a department of the German Federal Government. Three papers concern techniques for planning. One starts the development of a model of engineers' salaries to aid career development and professional utilization, followed by a model for capital–labour substitution decisions in the military field. The final paper briefly reviews the development and rationale for human resource accounting.

Finally, four authors provide short papers, including reviews of the conference and papers.

In sum, does the book live up to expectations? The answer must be 'only partially'. There are some papers here which will be referred to for many years, whereas others may have served to stimulate the conference but do not seriously increase the value of the text. It is not excessively expensive as books go today and it will be useful to workers in the field as well as being an essential library holding. But perhaps it is time to call a halt to the production of books of papers. There is no substitute for the coherent and well-presented text. Papers are more suitably enshrined in journals after a competent review procedure; even today the papers in this book which contribute to the science of human resources would probably have appeared more quickly in journals than via the medium of a hard-cover book.

E. N. CORLETT

Lithium Research and Therapy. Edited by F. N. Johnson. London: Academic Press. 1975. Pp. xxv+569. £13.50.

Currently, there is a rapid expansion in a wide range of fields in psychopharmacology and it is difficult, if not impossible, to keep abreast of events. Because of this rapid expansion in a particular field, a book which attempts to review the work is often out of date before it reaches the shelves of the bookseller. The editor of *Lithium Research and Therapy*, F. N. Johnson, is only too aware of this problem, and hopes that the contents will be used as a reference work and a statement of the position of lithium research and therapy at the time of going to press. In this way future trends can be evaluated against the present background.

The editor has assembled an impressive list of contributors and has divided the book into two major sections.

Part I is concerned with clinical studies and practical issues. Here the chemistry of lithium and how it can be used clinically is discussed, along with the use of lithium in a variety of psychiatric disorders, in particular manic depressive psychosis. Also in this part, chapters are devoted to the side effects and problems of patient management. This arrangement, in a sense, mirrors the development since 1949 of the use of lithium in that it was found almost by chance to have therapeutic value, and at first, the clinical advances outpaced the pharmacological.

Part II is concerned with physiological, biochemical and behavioural studies. Such topics as uptake, distribution and excretion of lithium are dealt with in depth, as is the effect of lithium on electrolyte balance, carbohydrate metabolism, acetylcholine and amine metabolism, and cyclic A.M.P. Of special interest to psychologists will be the two chapters written by the editor dealing with the behavioural effects of lithium. Here indeed is a tangled web of results. F. N. Johnson manages to organize them and bring some general trends together in a comprehensive and critical manner.

The editor succeeds admirably in his task of providing a comprehensive picture of the research background and clinical applications of lithium salts. Each chapter is clearly set out with relevant headings and references. It is a pity though that there is not an overall reference list.

As a reference book both for clinicians and for those working in multidisciplinary research projects this is a book of outstanding merit. It is both scholarly and practical. It should be on every postgraduate medical library shelf and indeed on every psychiatrist's if not every psychologist's as well.

ISABEL MOYES
DON KENDRICK

Br. J. Psychol. (1976), 67, 2, *pp.* 297–300
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

Arasteh, A. Z. *Toward Final Personality Integration.* Chichester: Wiley. 1975. Pp. xvii + 291. £7.40; paper, £4.00.

Argyle, M. & Cook, M. *Gaze and Mutual Gaze.* Cambridge: Cambridge University Press. 1975. Pp. xi + 210. £6.50.

Auerbach, S. & Rivaldo, J. A. (eds). *Child Care: A Comprehensive Guide.* New York: Human Sciences Press. 1975. Pp. xii + 251. $12.95.

Badcock, C. R. *Levi-Strauss. Structuralism and Sociological Theory.* London: Hutchinson. 1975. Pp. 125. £3.75; paper, £1.75.

Bloom, B. L. *Changing Patterns of Psychiatric Care.* New York: Human Sciences Press. 1975. Pp. 360. $15.95.

Barker, P. *Basic Child Psychiatry,* 2nd ed. London: Crosby Lockwood Staples. 1975. £6.00; paper, £3.60.

Booth, A. *Growing up in Society.* London: Methuen. 1975. Pp. 144. 75p.

Bower, F. M. (ed.). *Learning to Play.* New York: Human Sciences Press. 1975. Pp. 158. $4.98.

Bram, J. R. *Ancient Astrology Theory and Practice.* Bark Ridge, N.J.: Noyes Press. 1975. Pp. xi + 336. $15.00.

Brown, B. B. (ed.). *The Biofeedback Syllabus.* Springfield, Ill.: Thomas. $26.50; paper, $19.75.

Brown, H. & Stevens, R. (eds). *Social Behaviour and Experience.* London: Hodder & Stoughton. 1975. Pp. xii + 627. £8.50; paper, £4.50.

Bullmer, K. *The Art of Empathy.* New York: Human Sciences Press. 1975. Pp. vii + 160. $4.95.

Calfee, R. C. *Human Experimental Psychology.* London: Holt, Rinehart & Winston. 1975. Pp. xii + 532. £7.95.

Cattell, R. B. & Child, D. *Motivation and Dynamic Structure.* London: Holt, Rinehart & Winston. 1975. Pp. xi + 287. £6.50.

Cheshire, N. M. *The Nature of Psychodynamic Interpretation.* Chichester: Wiley. 1975. Pp. 229. £5.95.

Cofer, C. N. & Fitzgerald, H. E. (eds). *Psychology: A Programmed Modular Approach.* Homewood, Ill.: Learning Systems Ltd. 1975. Pp. vi + 674. $9.95.

Coles, R. *The Mind's Fate: Ways of Seeing Psychiatry and Psychoanalysis.* Boston, Toronto: Little, Brown. 1975. Pp. xvi + 282. $10.00.

Cornish, D. B. & Clarke, R. V. G. *Residential Treatment and its Effects on Delinquency.* London: HMSO. 1975. Pp. vi + 74. £1.00.

DePalma, D. J. & Foley, J. M. (eds). *Moral Development. Current Theory and Research.* Chichester: Wiley. 1975. Pp. viii + 206. £7.68.

Deutsch, M. & Hornstein, H. A. L. (eds). *Applying Social Psychology. Implications for Research, Practice and Training.* Chichester: Wiley. 1975. Pp. viii + 287. £8.00.

Dinkmeyer, D. & Carlson, J. *Consultation: A Book of Readings.* Chichester: Wiley. 1975. Pp. xvii + 295. £6.28.

Doob, L. W. *Pathways to People.* London: Yale University Press. 1975. Pp. x + 333. £13.75.

Dragastin, S. E. & Elder, G. H. *Adolescence in the Life Cycle.* Chichester: Wiley. 1975 Pp. viii + 324. £8.60.

Ehrlich, S. *Apprentissage et Memoire chez l'Homme*. Paris: Presses Universitaires de France. 1975. Pp. 338. No price given.

Evans, P. *Motivation*. London: Methuen. 1975. Pp. 143. 75p.

Flack, F. F. & Draghi, S. C. (eds). *The Nature and Treatment of Depression*. Chichester: Wiley. 1975. Pp. xii + 422. £11.25.

Fransella, F. *Need to Change?* London: Methuen. 1975. Pp. 144. 75p.

Freedman, A. M., Kaplan, H. I. & Sadock, B. J. *A Comprehensive Textbook of Psychiatry*, vols 1 and 2, 2nd ed. Baltimore: Williams & Wilkins. 1975. Pp. 2609. £65.00 both volumes.

Freedman, J. L. *Crowding and Behaviour*. San Francisco: W. H. Freeman. 1975. Pp. viii + 177. £2.50.

Golann, S. E. *Current and Future Trends in Community Psychology*. New York: Human Sciences Press. 1975. Pp. 235. $9.95.

Goldstein, M. J. & Palmer, J. D. *The Science of Anxiety*, 2nd ed. London: Oxford University Press. 1975. Pp. xviii + 294. £3.00.

Gove, W. R. (ed.). *The Labelling of Deviance*. Chichester: Wiley. 1975. Pp. 313. £9.25.

Graziano, A. M. (ed.) *Behaviour Therapy with Children*, vol. II. Chicago: Aldine. 1975. Pp. ix + 640. $19.95.

Greenberg, L. A. (ed.). *Psychodrama: Theory or Therapy*. London: Souvenir Press. 1975. Pp. xvi + 496. £4.50; paper, £2.75.

Gronwall, D. M. A. & Sampson, H. *The Psychological Effects of Concussion*. London: Oxford University Press. 1975. Pp. 118. £2.50.

Haber, R. N. & Fried, A. H. *An Introduction to Psychology*. London: Holt, Rinehart & Winston. 1975. Pp. 807. £5.25.

Hammerton, M. *Statistics for the Human Sciences*. London: Longmans. 1975. Pp. x + 182. £2.75.

Hammond, K. R. & Joyce, C. R. B. *Psychoactive Drugs and Social Judgement: Theory and Research*. Chichester: Wiley. 1975. Pp. xvii + 278. £9.25.

Howe, M. *Learning in Infants and Young Children*. London: Methuen. 1975. Pp. 216. £6.95; paper, £2.95.

Jessup, G. & Jessup. H. *Selection and Assessment at Work*. London: Methuen. 1975. Pp. 143. 75p.

Kagan, F., Harwood, T., Rickels, K., Rudzik, A. D. & Sober, H. *Hypnotics. Methods of Development and Evaluation*. Chichester: Wiley. 1975. Pp. 432. £16.00.

Kessen, W. (ed.). *Children in China*. London: Yale University Press. 1975. Pp. xiii + 241. £6.50; paper, £2.20.

Klein, M. *Love, Guilt and Reparation and Other Works* 1921–1945. London: Hogarth Press. 1975. Pp. xi + 468. £6.50; paper, £2.20.

Klein, M. *Envy and Gratitude and Other Works* 1946–1963. London: Hogarth Press. 1975. Pp. vii + 365. £6.00.

Koppitz, E. M. *The Bender Gestalt Test for Young Children*. Volume II. *Research and Application* 1963–1973. New York: Grune & Stratton. 1975. Pp. xiv + 205. £6.50.

Kris, E. *Selected Papers of Ernest Kris*. London: Yale University Press. 1975. Pp. x + 537. £11.00.

Levine, M. *A Cognitive Theory of Learning. Research on Hypothesis Testing*. Chichester: Wiley. 1975. Pp. xii + 313. £8.85.

Lindiall, C. M. & Nitko, A. J. *Measuring Pupil Achievement and Aptitude*, 2nd ed. New York: Harcourt, Brace Jovanovich. 1975. Pp. xi + 237. £2.75.

Lombardi, D. *Search for Significance*. Chicago: Nelson-Hale. 1975. Pp. xiii + 131. $6.95.

Lubin, A. J. *Stranger on the Earth. The Life of Vincent van Gogh*. St Albans: Granada. 1975. Pp. 304. £1.50.

MACCIA, E. S., COLEMAN, M. A., ESTEP, M. & SHIEL, T. M. (eds). *Women and Education.* Springfield, Ill.: Thomas. 1975. Pp. xiii + 301. $13.75; paper, $10.95.

MAHLER, M. S., PINE, F. & BERGMAN, A. *The Psychological Birth of the Human Infant.* London: Hutchinson. 1975. Pp. xii + 308. £6.71.

MANNINO, F. V., MACLENNAN, B. W. & SHORE, M. F. *The Practice of Mental Health Consultation.* Chichester: Wiley. 1975. Pp. x + 255. £8.00.

MEYER, B. J. F. *The Organization of Prose and its Effects on Memory.* Amsterdam: North Holland. 1975. Pp. xvii + 249. D. fl. 35.00.

MILLS, G. C., REISLER, R. L., ROBINSON, A. E. & VERMILYE, G. *Discussing Death.* Homewood, Ill.: ETC Publications. 1975. Pp. 140. $8.50; paper, $5.50.

MITTLER, P. *Research to Practice in the Field of Mental Handicap.* London: Institute for Research in Mental and Multiple Handicap. 1975. Pp. 49. 85 p.

MOHSIN, S. M. *Experiments in Psychology.* Bombay: Orient Longman. 1975. Pp. x + 340. Rupees 15.50.

MUELLER, R. J. *Practices of Classroom Learning and Perception.* London: George Allen & Unwin. 1975. Pp. xvii + 320. £6.00; paper, £2.95.

NAVACO, R. W. *Anger Control.* Farnborough: Lexington Books. 1975. Pp. xii + 134. £7.85.

PARKER, A. *States of Mind: ESP and Altered States of Consciousness.* London: Malaby Press. 1975. Pp. 198. £3.95.

PECK, D. & WHITLOW, D. *Approaches to Personality Theory.* London: Methuen. 1975. Pp. 144. 75p.

PHILLIPS, J. L. *The Origins of Intellect. Piaget's Theory*, 2nd ed. San Francisco: W. H. Freeman. 1975. Pp. xviii + 205. £4.40; paper, £2.00.

PINOL-DOURIEZ, M. *La Construction de l'Espace.* Paris: Delachaux et Mistle. Pp. 240. Price not given.

POLLACK, G. K. *Leadership of Discussion Groups. Case Material and Theory.* Chichester: Wiley. 1975. Pp. 391. £9.20.

PRUYSER, P. *Between Belief and Unbelief.* London: Sheldon Press. 1975. Pp. xviii + 301. £5.95.

RANDALL, J. L. *Parapsychology and the Nature of Life.* London: Souvenir Press. 1975. Pp. 252. £4.00.

RACHMAN, A. W. *Identity Group Therapy with Adolescents.* Springfield, Ill.: Thomas. 1975. Pp. ix + 323. $16.75.

RADFORD, J. & KIRBY, R. *The Person in Psychology.* London: Methuen. 1975. Pp. 143. 75p.

RANKIN, J. G. *Alcohol, Drugs and Brain Damage.* Toronto: Alcoholism and Drug Addiction Research Foundation of Ontario. 1975. Pp. xii + 101. $8.00; paper, $5.50.

RIEGEL, K. F. & ROSENWALD, G. C. (eds). *Structure and Transformation.* Chichester: Wiley. 1975. Pp. xviii + 250. £9.50.

ROBERTS, T. B. (ed.). *Four Psychologies Applied to Education.* Chichester: Wiley. £11.50; paper, £5.40.

ROSENHAN, D. & LONDON, P. *Theory and Research in Abnormal Psychology*, 2nd ed. London: Holt, Rinehart & Winston. Pp. xvii + 454. £3.75.

SCHULBERG, H. C. & BAKER, F. *Developments in Human Services*, vol. I. New York: Human Sciences Press. 1975. Pp. 404. $17.95; paper, $13.95; separates, $4.95.

SIEGAL, R. K. & WEST, L. Y. *Hallucination Behaviour, Experience and Theory.* Chichester: Wiley. 1975. Pp. xiii + 322. £14.00.

SINGER, R. N. *Motor Learning and Human Performance.* London: Collier Macmillan. 1975. Pp. xx + 849. £8.50.

TOBACCO RESEARCH COUNCIL. *Review of Activities*, 1970–1974. London: Tobacco Research Council. 1975. Pp. 114. Price not given.

VOLLE, F. C. *Mental Evaluation of the Disability Claimant.* Springfield, Ill.: Thomas. 1975. Pp. x + 121. $10.50.

WALL, W. D. *Constructive Education for Children.* London: Harrap; Paris: The UNESCO Press. 1975. Pp. xv + 349. £7.50.

WARBURTON, D. M. *Brain, Behaviour and Drugs.* Chichester: Wiley. 1975. Pp. x + 280. £6.95.

WEINBERG, G. *Know Yourself. Help Yourself.* London: Sheldon Press. 1975. Pp. 245. £4.95.

WHITE, R. J. *The Interpretation of Dreams.* Park Ridge, N.J.: Noyes Press. Pp. ix + 259. $15.00.

WHITEHEAD, J. M. *Personality and Learning*, vol. I. London: Hodder & Stoughton. 1975. Pp. vii + 407. £7.25; paper, £3.75.

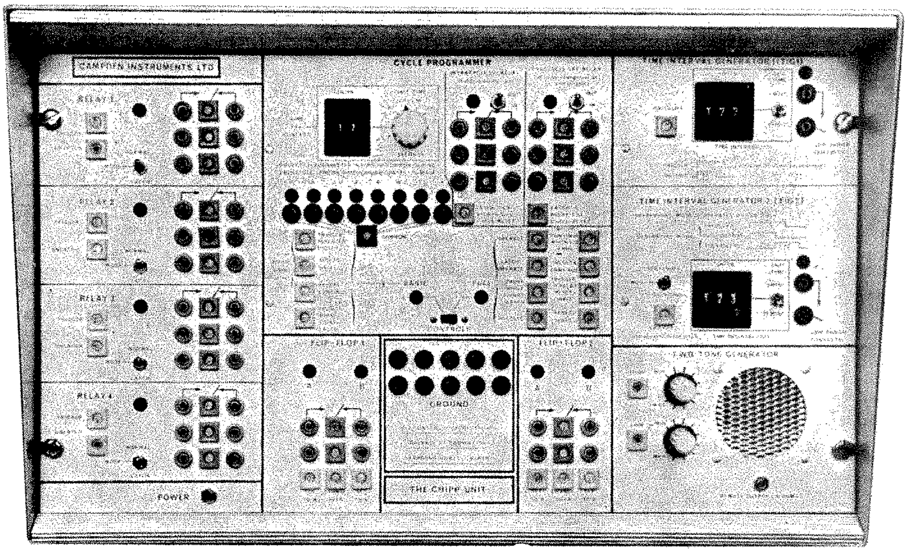
WHITFIELD, P. R. *Creativity and Industry.* Harmondsworth: Penguin Books. 1975. Pp. 217. 60p.

WOOTTON, A. *Dilemmas of Discourse. Controversies about the Sociological Interpretation of Language.* London: George Allen & Unwin. 1975. Pp. 125. £4.50.

ZELKIND, I. & SPRUG, J. *Time Research: 1172 Studies.* Meutchen, N.J.: The Scarecrow Press. 1975. Pp. v + 248. $8.50.

ZUSMAN, J. & BERTSCH, E. F. *The Future Role of the State Hospital.* London: Lexington Books. 1975. Pp. xviii + 410. £8.00.

NEW from CAMPDEN INSTRUMENTS – a powerful and unique package of control logic – THE CHIPP UNIT (Controller for Human Information Processing Psychology)



The CHIPP UNIT has been designed by CAMPDEN in consultation with Dr. John Frisby and Dr. Adrian Simpson of the Department of Psychology in the University of Sheffield who have also written the comprehensive Teaching Manual. The design of the CHIPP UNIT caters ideally for the present day trends in Psychological Experimentation thus superseding existing "Student Laboratory Units"

SPECIAL FEATURES INCLUDE:–

- LOGIC FACILITIES IN A SINGLE INSTRUMENT
- COMPLETE ELECTRICAL ISOLATION BETWEEN THE CHIPP UNIT AND OTHER EQUIPMENT
- L.E.D. INDICATOR LAMPS TO SHOW STATES AT ANY GIVEN INSTANT
- SIMPLE LOGICAL OPERATION BY GROUND SIGNALS
- CLEAR PANEL DESIGN FOR EASE OF OPERATION
- IDEALLY SUITED FOR STUDENT USE AS WELL AS RESEARCH – INCLUDING ANIMAL WORK

The CHIPP UNIT comes with a comprehensive USER'S MANUAL which includes a full description of how it is operated. This Manual includes illustrations on the programming of a number of psychological experiments and is written with the student user particularly in mind so that students should be able to teach themselves how to use it.

For more information on the CHIPP UNIT also on our FISH experimental units write to:–

CAMPDEN INSTRUMENTS LTD.,
186, CAMPDEN HILL ROAD,
LONDON, W8 7TH.

Telephones:– (01) 727 3437
(01) 229 3636

Explaining the Brain

W. Ritchie Russell and A. J. Dewar

How do we think? What is memory? What is pain? These questions are examined in a clear and simple account for the student and layman of how the human brain works. The book discusses nerve cells, and how they work together, the infant brain and its development, concentration and thinking, phantom limbs and psychosomatic disease, concussion, brain damage and the effect of drugs. Above all, readers will learn how they can help their children's brains to develop and how they should look after their own. £3.25 paper covers £1.25

Function and Evolution in Behaviour

Essays in honour of Professor Niko Tinbergen

Edited by Gerard Baerends, Colin Beer and Aubrey Manning

In this volume a number of well-known ethologists pay tribute to Professor Niko Tinbergen, one of the founders of ethology and a recent winner of the Nobel Prize for Physiology. The contributors are all former pupils or associates of his, and their essays discuss behavioural studies in a wide range of different species – from cephalopods to carnivores, from spiders to rooks. £16.50

Malnutrition and Brain Development

Myron Winick

A growing body of evidence shows that malnutrition early in life is associated with retarded brain development. This book reviews the experimental and clinical studies on the effects of both severe malnutrition – such as marasmus and kwshiorkor – and of milder forms of malnutritional deprivation. It will be of interest to nutritionists, psychologists, neurophysiologists, and epidemiologists. £5 *Oxford Medical Publications*

Race and I.Q.

Edited by Ashley Montagu

The aim of the articles reprinted in this book is to provide the reader interested in this difficult and controversial subject with the means of understanding the defects not only of certain claims made about the value of I.Q. tests but also the defects of such tests when applied to the measurement of 'racial' differences in intelligence. The author's conclusion is that it is unlikely that we shall ever be able to say to what extent differences in intelligence are due to genes or environment. £5.50 paper covers £2 *Galaxy Books*

Oxford University Press

Oxford Books for Students

Principles and Methods of Psychology

Robert B. Lawson, Steven G. Goldstein and Richard E. Musty

This book provides an introduction to the basic areas of psychology which include learning, motivation, perception, social psychology, individual differences, and experimental design and statistics. It is intended for use in experimental courses and in introductory courses where the emphasis is on experimental psychology. A teacher's manual with multiple choice test items is available on request from the University Promotion Section. Illustrated £8.50

Experiments in Psychology

Robert B. Lawson, Steven G. Goldstein and Richard E. Musty

The purpose of this laboratory manual and research kit is to provide the student with practical experience in conducting and reporting psychological experiments in the areas of perception, learning, motivation, and social psychology and individual differences. The manual contains some background information and procedural instructions to carry out the experiments in each of the above areas while the laboratory kit contains the requisite apparatus to complete each experiment. Illustrated £5.95

The Experience of Anxiety

A Casebook

Edited by Michael J. Goldstein and James O. Palmer

The second edition of this collection of case histories in abnormal psychology includes nine new cases covering phobia, hyperkinesia, death, suicide, female identity conflict, ethnic identity conflict, anorexia nervosa, drug abuse, and divorce. The new cases put more emphasis on the problems of young people and also on family interaction. Second edition paper covers £3

A Handbook for the Study of Suicide

Edited by Seymour Perlin

This book views the phenomenon of suicide from a historical, literary, and philosophical background. It then moves by way of the social sciences to a balanced psychiatric consideration of suicidal behaviour and presents the contribution that each relevant discipline makes to the study of suicide. It also offers a modern perspective on rational motivation toward suicide which diverges from the traditional psychiatric position that all suicides represent psychopathology. Illustrated £5.75 paper covers £3.45

Inspection copies of these books are available to lecturers wishing to consider them for adoption as required texts. Please write to University Promotion Section, Oxford University Press (BJP), Ely House, 37 Dover Street, London W1X 4AH, with details of the course and the number of students involved.

Brain

A Journal of Neurology

Volume 99 Part I March 1976

Peripheral Nerve Fibre Changes in Asymptomatic Children of Patients With Familial Amyloid Polyneuropathy. Jogre Carvalho, A. Coimbra and C. Andrade

Periodic Alternating Nystagmus. R. W. Baloh, V. Honrubia and H. R. Konrad

Progressive Encephalomyelitis with Rigidity – Its Relation to 'Subacute Myoclonic Spinal Neuronitis' and to the 'Stiff Man Syndrome'. A. M. Whitely, M. Swash and H. Urich

Delayed Visual Perception and Delayed Visual Evoked Potentials in the Spinal Form of Multiple Sclerosis and in Retrobulbar Neuritis. D. Regan, Beryl A. Milner and J. R. Heron

Aluminium, Neurofibrillary Degeneration and Alzheimer's Disease. D. R. Crapper, S. S. Krishnan and S. Quittkat

The Neuropathology of Progressive Rubella Panencephalitis of Late Onset. Jeannette J. Townsend, Jerry S. Wolinsky and J. Richard Baringer

Benign Myopathy, with Autosomal Dominant Inheritance – A Report on Three Pedigrees. Jaap Bethlem and George K. Van Wijngaarden

Effects of X-Radiation on the Spinal Cord: An Experimental Study of the Morphological Changes in Central Nerve Fibres. F. L. Mastaglia, W. I. McDonald, J. V. Watson and K. Yogendran

The Gate-Control Theory of Pain – A Critical Review. P. W. Nathan

Notices of Recent Publications

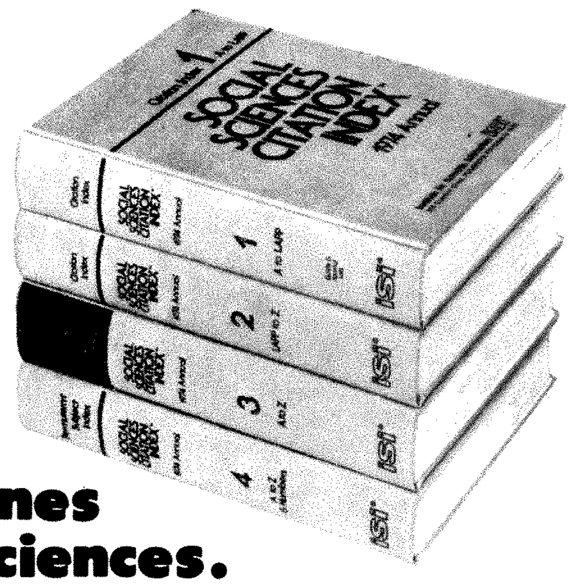
Subscription rate for an annual volume of four issues (March, June, September and December): £12.50 ($37) post free. Single issue: £3.50 ($10) net, postage extra. Orders and inquiries for back numbers should be addressed to: Wm. Dawson & Sons Ltd., Cannon House, Folkestone Kent.

Subscriptions to the current volume:

OXFORD UNIVERSITY PRESS
Press Road, Neasden, London NW10 0DD

Some indexes to the journal literature cover some of the social sciences disciplines . . . A few concentrate on one discipline, but

Only SSCI™ covers all the major disciplines in the social sciences.



The *Social Sciences Citation Index (SSCI)* provides the kind of coverage you need for doing in-depth searches of the journal literature. It covers every editorial item in over 1,400 leading social sciences journals in fields ranging from anthropology to urban planning. Plus it selectively indexes over 1,200 more journals from the natural, physical and biomedical sciences for articles relevant to the social sciences.

Because it is multidisciplinary, the *SSCI* does the work of several discipline-oriented indexes. A single look-up in the *SSCI* gives you access to the literature of over 50 major specialties. In all, nearly 85,000 items are indexed each year.

Easy to use, the *SSCI* offers several indexing approaches so you can quickly find pertinent information. All you need to start a search is an earlier paper on your topic, the name of a key author in the field, or simply a single term related to your subject. There are no specialized vocabularies to learn. No complex classification systems to master.

Fact is, with the *SSCI* in your library's reference section, you'll be able to locate articles reported in the journal literature of all the major disciplines of the social sciences more quickly and easily than ever before.

For detailed information on the *Social Sciences Citation Index*, complete the coupon below and mail it today.

Please send me more information on your *Social Sciences Citation Index*™.

Name ____________________

Position ____________________

Organization ____________________ Tel. ____________________

Address ____________________

City ____________________ State ____________________ ZIP ____________________

147-(2)446a

isi® **Institute for Scientific Information®**
325 Chestnut St., Phila., Pa. 19106, U.S.A., Tel: (215) 923-3300, Cable: SCINFO, Telex: 84-5305

©1976 ISI

The British Journal of Criminology

Delinquency and Deviant Social Behaviour

The purpose and policy of the Journal is to foster and publish researches and studies into the causes and prevention of crime, and the treatment of criminals.

Each number of the journal contains original papers. In addition the journal publishes general surveys of current developments in various countries. On occasion an individual part of the journal will consist of a symposium dealing with a special subject, with papers from groups of experts on issues not only purely scientific but of public moment at any given time.

There are in-depth notes and comments on practical developments, notably in the field of penal reform, on government or privately sponsored reports, and on new experiments in dealing with delinquent or deviant persons. Finally the Books and Periodicals section provides a comprehensive review of all significant literature in the field.

MAIN CONTENTS OF THE JANUARY 1976 ISSUE

Clifford R. Shaw and Henry McKay: Chicago Criminologists, by JON SNODGRASS.

Psychiatric Referrals from Edinburgh Courts, by MOYA WOODSIDE.

Psychological Correlates of Long-Term Imprisonment, by N. BOLTON, F. V. SMITH, K. J. HESKIN and P. A. BANISTER.

The Staging of Magistrate's Justice, by PAT CARLEN.

Twenty-Seven Robbers, by JOHN GUNN AND JEAN GRISTWOOD.

Illegal Abortions and the Abortion Act 1967, by PAUL CAVADINO.

1976 Subscription: £9.50 post free.

Published quarterly on behalf of The Institute for the Study and Treatment of Delinquency by

Stevens & Sons Ltd.

Subscriptions: North Way, Andover, Hants, SP10 5BE

British Journal of Educational Psychology

Vol. 46, Part 1 February 1976

The British Journal of Educational Psychology is concerned mainly with reporting the results of experimentally based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry.

The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words).

There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists.

Subscriptions (£6 a year, or £2 per issue) should be sent to: Scottish Academic Press Ltd., 33 Montgomery Street, Edinburgh EH7 5JX.

CONTENTS

SYMPOSIUM LEARNING PROCESSES AND STRATEGIES

EDITORIAL INTRODUCTION: The Verb "to Learn" Takes the Accusative

MARTON, F., AND SALJO, R. On Qualitative Differences in Learning: I – Outcome and Process

PASK, G.: Conversational Techniques in the Study and Practice of Education

ARTICLES

DONACHY, W.: Parent Participation in Pre-School Education

GALLOWAY, D.: Size of School, Socio-Economic Hardship, Suspension Rates and Persistent Unjustified Absence from School

CLIFFORD, H. M.: Are Common Sense Decisions Deterring Learning?

TAMIR, P.: The Relationship between Achievement in Biology and Cognitive Preference Styles of High School Students

BIGGS, J. B.: Dimensions of Study Behaviour: Another Look at ATI

RESEARCH NOTES

STANLEY, G.: The Processing of Digits by Children with Specific Reading Disability (Dyslexia)

SHAYER, M.: The Pendulum Problem

WEIR, C.: Interviews as an aid to Selection of Psychology Undergraduates

BOOK REVIEWS

PUBLICATIONS RECEIVED

SCOTTISH ACADEMIC PRESS LIMITED
33 MONTGOMERY STREET, EDINBURGH EH7 5JX

ARCHIVIO DI PSICOLOGIA
NEUROLOGIA E PSICHIATRIA
rivista trimestrale pubblicata a cura dell'Università Cattolica del Sacro Cuore
direttore Leonardo Ancona
comitato di redazione
Mario Bertini – Renzo Carli – Fabrizio Napolitani – Luigi Pizzamiglio
Corrado Pontalti – Carlo Saraceni – Ciro Sdino

YEAR XXXVI Nr. III July–September 1975

Leading article.

Cannizzaro, L. – Cecchini, M. – Musatti, T. Operation and code in child's operational evolution: proportion operation

Bergonzi, P. – Mazza, S. – Mennuni, G. – Zolo, P. Selective sleep deprivation (stage IV) in epileptic patients with partial seizures

Corradi Fiumara, G. Epistemophilia and symbolization process

Bassan, F. Psychanalysis and philosophy in Money-Kyrle's thought. II

Calegari, P. Methodological note on sequential researches in adolescents

Rossi, R. Seneca and Nero: inspirer and killer in a matricide. A psychanalytical standpoint

Quadrio, A. – Sacchi Casarotti, M. Conflictual attitudes towards the ecological problem

NOTES AND DISCUSSIONS

Venini, L. New meeting opportunities between psychanalysis and literature

Every article is followed by an Italian, French, English and German summary.

Every issue summarizes many Italian and Foreign works of psychology, neurology and psychiatrics and gives large review of news and information.

Redaction – Administration: 1, Largo A. Gemelli – 20123 MILANO (Italy).
Annual subscription: Italy 11.000 L.; foreign countries L. 16.000 (or equivalent foreign money).

REVUE DE PSYCHOLOGIE APPLIQUEE

Publication Trimestrielle

Directeurs: Dr P. Pichot et P. Rennes

4ème Trimestre 1975 – Volume 25 – No. 4

SOMMAIRE

M. Auzias – Une épreuve de latéralité usuelle et rapports entre latéralité usuelle et latéralité graphique chez l'enfant de 5 à 11 ans.

L. Lurçat – Répétition et alternance dans la conservation d'un modèle graphique.

J. Poitrenaud et C. Moreaux – Réponses données au test de Rorschach par un groupe de sujets âgés cliniquement normaux. Analyse quantitative des protocoles en fonction de l'âge.

Edwin A. Fleischman – Portrait, notice biographique et bibliographique.

Rédaction et Administration:
48 Avenue Victor Hugo, 75783 PARIS CEDEX 16

Abonnements (un an) France: 45 F; Étranger: 55 F
Le numéro France: 12 F; Étranger: 12 F

Règlements au nom des Éditions du Centre de Psychologie Appliquée

Numéro spécimen sur demande

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
The Journal of Occupational Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *The Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1976 subscription prices, payable in advance are: £20.00 net per volume post free (US $60.00 in the USA and Canada) for *The British Journal of Psychology*; £16.00 net (US $48.00) for *The British Journal of Medical Psychology*; £13.00 net (US $39.00) for *The Journal of Occupational Psychology*; and £19.00 net (US $57.00) for *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian inquiries should be addressed to the Cambridge University Press, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15.00 per volume post free ($34.00 in the USA). Subscriptions should be sent to the Society's Agents, John Wright & Sons, Ltd., 42–44 Triangle West, Bristol BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £6.00 a year post free, payable in advance, or £2.00 each number. The subscription price in the USA is $20.00. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. D. B. Clarke, Department of Psychology, The University, Hull HU6 7RX;

for the *British Journal of Medical Psychology* should be sent to

Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London SW20 0NE;

for *The Journal of Occupational Psychology* should be sent to

Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster LA1 4YF;

for *The British Journal of Social and Clinical Psychology* should be sent to

Dr Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ: or to
Dr H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester M20 8LR;

for *The British Journal of Educational Psychology* should be sent to

Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable price 50p net (USA $2.00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18.00 (£16.00 if paid by Banker's Order). They are entitled to purchase any or all of the Society's Journals at special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London W1X 4DN.

The British Journal of Psychology *Volume 67, Part 2, May 1976*

CONTENTS

BOOTH, D. A., LEE, M. AND MCALEAVEY, CHRISTINE
Acquired sensory control of satiation in man *page* 137

KEMPEN, GERARD
Syntactic construction as retrieval plans 149

MORRIS, P. E. AND RIDGWAY, J.
Recency in free recall: STM plus LTM or STM alone? 161

NEWSTEAD, STEPHEN E.
Semantic constraints and sentence perception 165

BERLYNE, D. E. AND DITKOFSKY, JOYCE
Effects of novelty and oddity on visual selective attention 175

DAVIES, PETER
Conditioning afterimages: a procedure minimizing the extinction effect of normal test trials 181

HAMILTON, VERNON AND LAUNAY, GILLES
The role of information-processing capacity in the conserving operation 191

JAHODA, GUSTAV
Reproduction of Kohs-type figures by Ghanaian children: orientation errors revisited 203

HOWARTH, EDGAR
Were Cattell's 'personality sphere' factors correctly identified in the first instance? 213

WETHERICK, N. E. AND DOMINOWSKI, R. L.
How representative are concept attainment experiments? 231

MCASKIE, MICHAEL AND CLARKE, ANN M.
Parent–offspring resemblances in intelligence: theories and evidence 243

BOOK REVIEWS 275

OTHER PUBLICATIONS RECEIVED 297

© The British Psychological Society 1976

10 DEC 1976

Printed in Great Britain at the University Printing House, Cambridge

Volume 67, Part 3 *August 1976*

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by A. D. B. CLARKE
with J. BROWN, R. CONRAD, N. F. DIXON,
G. W. GRANGER, I. MARTIN, R. L. REID,
B. TIZARD *and* P. B. WARR

UNIVERSITY OF CALCUTTA

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET NEW YORK, N.Y. 10022

Assistant Editor: D. I. WILLIAMS
Book Review Editor: D. C. KENDRICK
Press Editor: MICHELE BENJAMIN

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements.

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of articles published from Vol. 56 (1965). A brief summary, suitable for use as an abstract should precede the article.

Manuscripts should be submitted in duplicate

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the author's names and corresponding to an alphabetical list at the end of the paper. This list must give author's names with all the details and the layout illustrated in the following examples:

Allport, G. W. & Vernon, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–284.

Galton, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

Révész, G. (1951). Le problème du génie. *Anée psychol.* **50**, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

© The British Psychological Society 1976

Permissions

For permission to reproduce material from *The British Journal of Psychology*, please apply to the London or New York office of Cambridge University Press.

ISI Tear Service, 325 Chestnut Street, Philadelphia, Pennsylvania 19106, USA is authorized to supply single copies of separate articles for private use only.

Subscriptions

The British Journal of Psychology is published quarterly.

The subscription price (including postage) of volume 67, 1976, is £20.00 net (US $60.00 in the USA and Canada). Single parts cost £6.00 each (US $18.00 in the USA and Canada).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1976), 67, 3, *pp.* 301–314
Printed in Great Britain

THOUGHT AND THE BRAIN*

By O. L. ZANGWILL
Cambridge Psychological Laboratory

I regret that not even the title of this Address is original. I have borrowed it from that of an old book by Henri Piéron called *Le Cerveau et la Pensée* (1923). In this book, the author's aim was to link, as and where possible, such knowledge as we have in psychology with what is known about the structure and functions of the brain. To his task, Piéron brought much erudition and a biological outlook not always so clearly in evidence today. In particular, he thought proper to adduce evidence not only from anatomy and physiology but also from clinical neurology, a field of medicine only too seldom represented in contemporary psychological texts. If I have purloined his title, it is because my own work over the years owes much to Piéron's example and reflects in considerable measure his own approach to psychological issues.

Lest some of you may think that I am a 'reductionist', as contemporary idiom has it, i.e. that my aim is to explain – or explain away – psychological phenomena in terms of physics and chemistry, I must hasten to say that nothing is further from my intention. You will hear nothing from me about nervous impulses or receptive fields, about transmitter substances or molecular models of memory, important as all of these may ultimately prove for explanation in psychology. My interest lies wholly within the traditional field of psychology, in human experience and human behaviour, and more especially in the ways in which study of the effects of brain injury may throw light upon their nature. Indeed the only real difference between my approach and that of those, such as John Cohen, who advocate a specifically humanistic psychology (Cohen, 1970), is that I prefer to think of man's humanity as vested in his bodily structure and function, not least, in the activities of his brain.

I. Fechner and psychophysics

As a Professor of Experimental Psychology, it is perhaps only proper that I should begin by paying tribute to the man who is generally accepted as the founder of my discipline. I refer of course to Gustav Theodor Fechner, the 'father of experimental psychology', whose famous book *Elemente der Psychophysik* appeared in 1860. Fechner is remembered today almost exclusively for his psychophysical law, which inspired what William James so slightingly dismissed as 'that dreadful literature' (James, 1890, 1, p. 549). But James had in mind only what Fechner himself termed *outer psychophysics*, namely, the correlation of sensory magnitude with the intensity of physical stimulus. But this, Fechner thought, was only the first step. Much more important, he believed, was what he called *inner psychophysics*, the correlation of subjective magnitude with the intensity of the central excitatory process. He thought that inner psychophysics would ultimately make possible a quantitative treatment not only of sensation but also of images, affects and indeed states of consciousness generally. Although techniques appropriate to such study were not available in Fechner's time and are indeed only in their infancy today, it is at least clear that, for him, psychophysics would eventually evolve into psychophysiology and that the whole idea of experimental psychology implied a close and essential concern for the cerebral basis of mind.

---

* Presidential Address to the British Psychological Society delivered at Nottingham on 5 April 1975.

Although Fechner did not take his idea of inner psychophysics very far, much of what he wrote was strikingly in advance of its time. For example, he made much play with the notion of a limen of consciousness, which might be affected both by sleep and wakefulness as well as by shifts of attention, voluntary or involuntary. This notion has been adequately expounded by Boring (1929) and I shall not consider it further here. I would however like to say something about Fechner's general ideas on consciousness which as I have suggested elsewhere (Zangwill, 1974) have surprising relevance to modern work.

Much influenced by the French physiologist Flourens, and anticipating Lashley over half a century later, Fechner placed great stress on the equipotentiality of the cerebral cortex. Consciousness he regarded essentially as an attribute of the cerebral hemispheres, although insisting that not all parts of them are of equivalent importance for mental life. Above all, he contended that continuity of anatomical structure is an essential condition of the unity of consciousness, a thesis later developed by the German philosopher, Eduard von Hartmann (1931). This led Fechner to consider closely the issues raised by the brain's duplex structure. If, he argued, it were possible to divide the brain longitudinally in the mid-line – an experiment which, incidentally, he considered impossible to achieve – something like the duplication of a human being would be brought about. 'The two cerebral hemispheres', he wrote, 'while beginning with the same moods, predispositions, knowledge and memories, indeed the same consciousness generally, will thereafter develop differently according to the external relations into which each will enter' (1860, **2**, p. 537). In short, splitting the brain will divide the stream of consciousness.

Hardly surprisingly, Fechner's position found its opponents, among them a once famous founder-member of this Society, William McDougall. In his *Body and Mind* (1911), McDougall came out strongly against the view that the unity of consciousness is conditional upon the continuity of nervous structure. In this connexion, the late Sir Cyril Burt has related* that McDougall more than once tried to bargain with the great Sherrington that if ever he (McDougall) should be smitten by an incurable disease, Sherrington should cut through his corpus callosum, that massive band of nerve fibres uniting the two hemispheres. 'If the physiologists are right', said McDougall, 'the result should be a split personality. If I am right, my consciousness should remain a unitary consciousness.' And this, Burt comments, he seemed to regard as the most convincing evidence of something like a soul!

---

* In a private letter to the late Professor R. C. Oldfield, dated June 29 1949. I am indebted to the Lady Kathleen Oldfield for her kind permission to reproduce this extract.

II. THE SPLIT-BRAIN EXPERIMENT

As we now know, the experiment thought by Fechner to be impossible and by McDougall to be practicable only in the case of a scientific martyr, is now a recognized neurosurgical procedure. I should, however, make clear that this drastic operation is carried out only in cases of severe and intractable epilepsy that have failed to respond to anticonvulsant medication. Known as commissurotomy, the operation was introduced by van Wagenen and Akelaitis in the early 'forties and revived in the 'sixties by Philip Vogel and Joseph Bogen, of the California College of Medicine. Roger Sperry, a distinguished psychologist and the first to develop experimental split-brain surgery in animals, was quick to see the gain to science from careful psychological study of individuals who had undergone this operation. At his kind invitation, I was privileged to spend some time in his Department at the California Institute of Technology five years ago and to gain first-hand experience of split-brain patients and their investigation.

I would like to say a little about the work of Sperry and his associates, which I regard as of first-rate importance. What are the effects of brain bisection upon psychological capacity? Does the operation really produce a doubling of the human being, as prophesied by Fechner? Or does consciousness remain unitary, as McDougall contended? At first blush, the advantage might appear definitely to lie with McDougall. Neither Akelaitis, nor after him Bogen and Sperry, could detect any gross change in intellect, personality or general behaviour of the split-brain patient after recovery from the immediate effects of the operation. So in spite of brain bisection, it would seem that the person remains unitary.

But this is not the whole story. Using techniques of a simplicity bordering upon the brilliant, Sperry and his co-workers have been able to show beyond reasonable doubt that each hemisphere of the brain may under appropriate circumstances operate independently of its fellow (see, for example, Sperry, Gazzaniga & Bogen, 1969). For example, if the eyes are closed an object actively explored by the left hand is not recognized by the right hand and vice versa. And similarly, the picture of an object exposed in the left half-field of vision is not recognized when re-exposed in the right half-field. These perceptual and discriminative tasks are carried out unilaterally and there is no apparent transfer of information between the hemispheres. There is even evidence that the two hemispheres may carry out concurrent tasks, such as copying designs flashed simultaneously to the two visual half-fields, without mutual awareness of each other's activity. Under conditions of unilateral input, therefore, each hemisphere functions as an independent processor, producing results reminiscent of the behaviour of two separate individuals. In Sperry's words, 'each hemisphere . . . seems to have its own conscious sphere for sensation, perception and other mental activities and the whole realm of gnostic activity of the one is cut off from the corresponding experience of the other hemsiphere' (Sperry *et al.* 1969).

Thus Fechner, too, has been proved right: Duplex conscious activity of the two hemispheres can result from callosal section. One must, however, remember that Fechner was writing before Broca and others had established that the two hemispheres appear to differ with regard to the representation of language. This asymmetry in the control of speech, originally established on the basis of clinical study, is beautifully illustrated in Sperry's work (Gazzaniga & Sperry, 1967). It is found that objects explored by active touch, or pictures or words exposed in the visual half-fields, can be named or described only when explored by the right hand or exposed in the right half-field. That is to say, expressive speech appears to be the prerogative of the left cerebral hemisphere. Although the right hemisphere does appear to possess limited powers of comprehension, its expressive capacities are extraordinarily limited. This is probably due to the left hemisphere possessing a virtual monopoly in control of the motor systems involved in linguistic expression, whether by speech or writing.

It has been argued by some critics, most notably Sir John Eccles (1965, 1973), that duplex processing does not necessarily imply duplex consciousness. Eccles has repeatedly stated that he regards the right hemisphere as a computer and its activities as devoid of consciousness. He argues first, that consciousness cannot be assumed in the absence of some form of symbolic communication; and secondly, that there is no evidence to suggest that the right hemisphere can, in and by itself, initiate voluntary activity. I have taken issue with Eccles elsewhere (Zangwill, 1974) and do not wish to repeat my arguments here. I may, however, perhaps remind you that it is possible for an adult to survive the loss of his entire left hemisphere without any apparent change in his level of consciousness. A patient that I was once able to study by courtesy of Dr Aaron Smith (1966, 1972), whose left hemisphere had been totally excised on account of an extensive and fast-growing cerebral tumour, was alert, cooperative, and fully responsive to all around him. In spite of severe aphasia, he could understand a good deal of what was said to him and was even able to sing a little! Emotional frustration was forcibly expressed by constant bad language. This patient was quite able to initiate voluntary movement of the non-paralysed parts of his body and even able to achieve a low-average score on a diagrammatic intelligence test. In view of facts of this kind, as well as the evidence assembled by Sperry and his co-workers as to the concurrence of unrelated perceptual events in the two disconnected hemispheres (Levy, Trevarthen & Sperry, 1972), I think we can accept Fechner's argument that consciousness is linked with the waking activity of either or both cerebral hemispheres and that its unity normally depends upon the structural link between them. Eccles' argument, like that of McDougall, is little more than a desperate rearguard action to safeguard the existence and indivisibility of the soul.

III. CEREBRAL DOMINANCE

I would now like to consider in rather more detail the implications of asymmetry in cerebral hemisphere function for issues in general psychology, with special reference to the differential effects of unilateral cerebral lesions. Traditionally, it has been supposed that one hemisphere, as a rule the left but occasionally the right, 'takes the lead' in the acquisition of language and manipulative skill. In consequence, aphasia and kindred disorders of speech are widely thought to result from lesions of the hemisphere contralateral to the preferred hand. This general relation between speech and handedness is crystallized in the phrase *cerebral dominance*.

A good deal of my own work over the years has been concerned with this question of cerebral dominance and its relation to handedness (Zangwill, 1960*a*, 1964*a*).

While the correlation between right-handedness and left-brainedness (i.e. aphasia resulting from lesions of the left hemisphere) has not been seriously questioned, my own work and that of many others has cast considerable doubt upon the complementary correlation between left-handedness and right-brainedness (cf. Humphrey & Zangwill, 1952; Ettlinger, Jackson & Zangwill, 1956; Zangwill, 1960*a*). Indeed a majority of left-handers or those with mixed patterns of lateral preferences appear to develop aphasia from lesions not of the right hemisphere but of the left hemisphere. None the less, aphasia from lesions of the right hemisphere does appear to be more common among left-handers than among right-handers, suggesting that there is indeed some relation, if an obscure one, between handedness and the acquisition of speech. We are at present collecting information about aphasia from right-sided lesions as it presents in both right- and left-handed individuals which it is hoped may throw further light on this vexed problem.

Apart from handedness, however, there does seem to be evidence that the left hemisphere has particular relevance to the initiation of voluntary activity. Although it would be naive to try to locate the Will in either hemisphere, there seems little doubt that purposive motor activity proceeds for the most part under left hemisphere control. Indeed this is probably the reason why the general behaviour of the split-brain patient differs so little from that of the normal individual. It is also noteworthy that difficulty in demonstrating the use of common objects, which neurologists describe as *ideational apraxia* (Zangwill, 1960*b*), appears to be a syndrome of the dominant cerebral hemisphere. It may indeed be that cerebral dominance evolved in the first place in connexion with tool-using and manipulative skill and only secondarily in relation to the evolution of language.

Until recently, it was widely accepted that, apart from language, equivalent lesions of the two hemispheres produce identical patterns of disability. Here again, some of my own work has suggested that lesions of the right, or minor, hemisphere may give rise to psychological deficits rather different from those associated with comparable lesions of the dominant hemisphere. Paterson and I first stumbled on this possibility in our work on brain injuries during the last war (Paterson & Zangwill, 1944), though we did not at first perceive its full implications. By and large, posterior lesions of the right hemisphere are prone to give rise to striking disorders in visuo-spatial perception and in the execution of constructional tasks under visual control. These are both more marked, and qualitatively different from, somewhat similar perceptual deficits arising from lesions of the left hemisphere. In the first place, neglect of the left-hand side of visual space is often extremely prominent. And in the second place, there is a noteworthy failure in the analysis of spatial relationships, leading to gross breakdown on simple tests of copying, drawing and assembling mosaics (McFie, Piercy & Zangwill, 1950; Ettlinger, Warrington & Zangwill, 1956). In the case of left hemisphere lesions, on the other hand, difficulties on such tasks tend to arise from deficits in form or colour perception, from loss of knowledge or manual skill, and may involve a general defect in abstraction (McFie & Zangwill, 1960).

These observations have been greatly strengthened by the outcome of split-brain studies. Bogen, in particular, has reported that, at all events in the earlier post-operative stages, the patient cannot write with his left hand, as might indeed be expected, but neither can he draw nor copy with his right. For example, even though right-handed, the patient makes a much better shot at drawing a cube with his left hand than with his right (Bogen & Gazzaniga, 1965). Further, careful experiments by Levy *et al.* (1972) have shown that if displays such as photographs of human faces or other complicated designs difficult to specify in words are shown to both hemispheres simultaneously and the patient required to respond by pointing to its duplicate in a multiple-choice array, the split-brain subject shows a marked predilection to react to the right hemisphere input. It is only if a verbal response is demanded that there is a shift to left hemisphere control. Moreover, the efficiency of visual discrimination is greater when input is processed by the right hemisphere than by the left. Unless verbal coding is essential, therefore, it seems likely that the right hemisphere typically takes the lead in processing complex visual information.

Even more interesting, perhaps, is the suggestion by Levy-Agresti & Sperry (1968) that the two hemispheres exhibit a characteristic difference in approach to visual tasks. In the case of the left hemisphere, the subject proceeds by sequential analysis of key features and makes obvious use of verbal coding. In the case of the right hemisphere, on the other hand, he proceeds by global impression displaying what used to be called synthetic rather than analytical perception. I suggested in my Sir Frederic Bartlett lecture three years ago (Zangwill, 1972) that these differences in strategy parallel closely those drawn by Bartlett (1932) between the vocalizer and the visualizer. Bartlett pointed out that, in remembering, vocalizers tend to rely largely on verbal analysis and are better than visualizers at retaining an order of presentation. Visualizers, on the other hand, tend to recall in more direct and pictorial a manner and make little use of verbal coding. From the point of view of accuracy, fortunately, neither method seems greatly superior to the other.

Vision is not the only sense which appears to bear a special relation to the activities of the right hemisphere. In her work at the Montreal Neurological Institute, Brenda Milner (1962) has reported significantly greater impairment in certain psychoacoustic tasks, in particular the perception of tonal quality and pattern, in patients with right-sided than in those with left-sided temporal lobe excisions. In trained musicians, too, there is evidence to suggest that disorders of musical appreciation and performance occur surprisingly often with right-sided cerebral lesions. It seems possible therefore that the right hemisphere has a special role to play in a number of human accomplishments that do not call for linguistic sophistication.

In this brief and summary account of interhemispheric relations and cerebral dominance, I have limited myself almost exclusively to clinical studies as it is in these that my experience principally lies. I need not remind you, however, of the many and ingenious attempts that have been made in recent years to study these issues using experimental techniques such as dichotic listening and the differential measurement of perceptual efficiency in the right and left half-fields of vision (cf. Blakemore, Iversen & Zangwill, 1972). These studies have in general borne out the findings of clinical observation and are helping us to understand their nature; not least, the role of each hemisphere respectively in directing and sustaining attention. Mention, too, should be made of attempts to study hemispheric asymmetries by means of physiological techniques, such as electroencephalography. In spite of the many hazards that enter into such experiments and the wide range of individual differences in results, we see in such work a promise that Fechner's inner psychophysics may yet become a reality.

IV. THOUGHT AND LANGUAGE

So far, we have considered relatively low-level psychological performance, e.g. perception, discrimination and motor skill, in relation to the duplex organization of the brain. But what of the higher mental processes? What of intelligence, problem-solving and thought? Here, too, I think an approach from the neurological standpoint is not without its promise. Let us consider first the activities of the left hemisphere which, as we have seen, are closely related to the acquisition and use of language and which many people have supposed to bear an important relation to conceptual thought.

In the study of aphasia, there has been a good deal of disagreement as to whether intellectual capacity is or is not primarily affected. As we know, most aphasics are capable of exercising adequate judgement in practical matters, and, as Weisenberg & McBride (1935) were the first to demonstrate, intelligence as measured by appropriate non-verbal methods may remain at a relatively high level. None the less, I do not think it possible to maintain that in aphasia only language is impaired; thinking, too, is appreciably affected. For example, aphasic patients are apt to experience great difficulty in what I have called the understanding and explanation of overall meaning (Zangwill, 1964*b*). If, for example, a patient whose speech disorder is only minimal is asked to explain, say, a simple proverb, he will frequently repeat it over and over again and evidently accepts it as entirely familiar. Yet he finds himself totally unable to find other words in which to express its meaning. Other patients, particularly those with left frontal lobe lesions, find difficulty in initiating conversation and seem to lack verbal fluency, not only in retrieval of words but of ideas as well. This Alexander Luria (1968) has described as 'dynamic aphasia'. I have noticed that such patients often do very badly on Liam Hudson's tests of divergent thinking (Hudson, 1966), suggesting that the patient's thought has become more stereotyped, less productive and less original. Such limitations in thought are seldom seen with lesions of the right hemisphere, suggesting that the left is specialized not only for language in the narrower sense, but also for those thought processes which find expression in linguistic form (Zangwill, 1966).

This conclusion is borne out, to some extent at least, by studies of recovery from aphasia in intelligent and well-educated subjects. Elvin & Oldfield (1951), for example, charted the progress of a university history student rendered dysphasic by a war wound of the brain. They analysed in great detail a series of tutorial essays covering the whole period of this student's university career. By and large, consistent progress was made in both form and content. First, words, and in particular prepositions, were increasingly less often misused. Secondly, 'over-general' words, e.g. 'things', gradually dropped out and expression became progressively more precise. And thirdly, the capacity to express ideas clearly and logically improved continuously. Whereas at first ideas were often 'telescoped' and the subject was apt unwittingly to change his viewpoint in expounding his argument, his later essays were well organized in content and satisfactory in style. Although it is difficult to disentangle the effects of education from those of natural recovery, it certainly appeared that in this young man improvement in formal composition proceeded *pari passu* with improvement in the organization and logical structure of his thinking.

Not surprisingly, the effects of aphasia on creative writing are catastrophic. Baudelaire is a characteristic example. Whereas several painters and even a few composers have been able to continue their work in spite of strokes causing serious damage to the left hemisphere, I know of no case of a writer similarly fortunate. Indeed, at a more mundane level, the after-effects of aphasia are always more serious for people such as secretaries and journalists than for those whose occupations are primarily technical. The indispensable instrument of the writer is not so much the pen as the left cerebral hemisphere.

V. VISUAL THINKING

What of the right cerebral hemisphere? I have already pointed out that lesions of its posterior portions are apt to produce marked defects in visuo-motor skill. In attenuated form, such defects may persist indefinitely, as Freda Newcombe (1969) has shown in a careful follow-up study of patients with circumscribed war wounds of the brain. But right hemisphere lesions may do more than impair perception; they may, if less constantly, produce appreciable handicap in visual recognition, particularly of faces and places, and in visual memory. Some patients report that their visual imagery has become faint and sketchy and others remark that they have ceased to dream (Humphrey & Zangwill, 1951). Even the illusion of *déjà vu* seems to bear a statistically significant relation to epileptogenic lesions of the right temporal lobe (Cole & Zangwill, 1963).

There is thus *prima facie* evidence that the right hemisphere has an important part to play in what might loosely be called visual thinking. The impoverishment of visual imagery and the cessation of dreaming suggest, perhaps, a deficiency in the kind of thinking commonly described as imagination or fantasy and regarded by Freud as the primary process of thought, closely related to unconscious activity. But visual thinking is also involved in many high-level conscious activities, as, for example, the solution of diagrammatic intelligence tests. Although Newcombe was unable to ascertain any significant difference in score on Raven's Matrices as between men who had sustained unilateral lesions of either the right or the left hemisphere, many other studies have given evidence of a significant inferiority of the right-hemisphere cases (Piercy & Smyth, 1962). It is also of some interest to note that, in Aaron Smith's studies of hemispherectomy, all three of his patients with right-sided hemisphere excision performed excessively poorly on this test. Only the patient with a left hemisphere excision managed to achieve a score within the low average range.

At the risk of being somewhat speculative, I have often been led to wonder whether, in the kind of problem-solving which proceeds by what used to be called insight, the right hemisphere may not play a predominant role. This kind of thinking has no need of language, as Köhler (1925) so brilliantly demonstrated in his early work on the mentality of apes. As you will remember, the problems he set his chimpanzees were essentially visual problems, using tools and overcoming obstacles, in which what he described as a restructuring of the field was essential to their solution. In spite of Washoe and other recent stars of the primate firmament, would it be wholly facile to suppose that the chimpanzee has in effect two right hemispheres?

In man, the development of propositional language renders the kind of thinking exemplified by Köhler's apes difficult to demonstrate, except perhaps in the split-brain patient. None the less, insight is not a bad way of describing what goes on when one suddenly sees the point of a humorous drawing or political cartoon. It also describes well the solution of puzzle-pictures, such as the 'Horse-and-Rider' problem used by Bartlett (1957) as an example of one-step problem-solving.

In this connexion, I may add that Bartlett himself called attention to the somewhat idiosyncratic thinking of creative artists, apparently so different from the 'closed system' variety of thinking characteristic of scientists and logicians and so favourite a subject of study in the psychological laboratory. 'There have indeed', wrote Bartlett, 'been persons who apparently wish to maintain that certain sorts of artists, especially those whose proper medium is colour, forms, shape and tone – painters and musicians – do not think at all' (1957, p. 187). Perhaps it would be fairer to say that artists and musicians think, but that their thinking differs in important respects from that of the scientist or engineer.

Can neurology throw any light on this matter? As I have said, the painter, unlike the writer or poet, is often able to maintain his work at a high level in spite of severe aphasia, often complicated by an enforced change of handedness. This might suggest that the right hemisphere is critically implicated in artistic endeavour. On the other hand, the effect of a right hemisphere lesion on the work of professional artists does not appear anything like as crippling as one might have anticipated. Professor Richard Jung, of Freiburg, who has made a special study of painters who have sustained right-sided strokes, notes only a striking and on occasion persistent neglect of the right half of the canvas – that is to say, of the part of the painting to the artist's left. There may also be certain changes in technique and manner of execution, probably in part compensatory (cf. Gardner, 1975). It would seem, therefore, that the right hemisphere is by no means as vital to artistic expression as is the left to literary creation.

One recent author, Robert Ornstein (1972), has made large claims on behalf of the right hemisphere as the seat of intuition as opposed to analytical thought. He claims that Western culture, with its stress on logic and literacy, has done much to damp down a whole world of intuitive and mystical experience.

While I am fully prepared to admit, with Ornstein, that the disconnected right hemisphere of the split-brain patient reacts to visual input in a global, quasi-intuitive manner, and even that the right hemisphere of a normal individual may well be the leading hemisphere in visual and spatial thinking, I can see no reason to link the activities of this hemisphere in any exclusive way with the manifestations of transcendental meditation or other altered states of consciousness. While it is of course possible that the right hemisphere is specially concerned in such activities, it is equally concerned in such typical manifestations of Western culture as the reading of maps and circuit diagrams, the design of architectural plans and the lay-out of aircraft instrument panels. Let us, therefore, be wary of an updated phrenology that seeks to provide a scientific justification for some, to me at least, irrational and disturbing trends in modern throught. Albeit mute, the right hemisphere is neither less nor more rational than its articulate fellow.

VI. DEVELOPMENTAL DISORDERS

The differences between the effects of right and left hemisphere lesions in the adult find a certain parallel in the developmental handicaps not uncommonly encountered in schoolchildren. On the one hand, we see children with marked and often rather selective backwardness in learning to read and spell. Such children, who are often referred to as dyslexic, have as a rule been rather slow in learning to talk and tend to score very much more poorly on verbal than on non-verbal intelligence tests. On the other hand, we have children whose level of verbal ability is consistently superior to the non-verbal and who show marked disability in spatial perception and manipulative skill. One inquiry which we undertook in Cambridge schools some years ago (Brenner, Gillman, Farrell & Zangwill, 1967), and in which over 800 children were tested, disclosed that roughly 3 per cent of the children gave evidence of a discrepancy of at least two S.D.s between their scores on a visual–spatial test battery and on a standard verbal intelligence test. Such children are typically awkward and clumsy and often in trouble on this account. Their mothers tell one that they were slow to learn to do up buttons or tie their shoelaces. At school, their handwriting is badly formed and uneven and their sums poorly set out, leading to frequent mistakes in simple arithmetic. They draw badly and are abysmally poor at such activities as carpentry or needlework. Worst of all for schoolboy prestige, they are singularly inept at ball games.

I do not wish to suggest that these patterns necessarily result from brain injury, though in some cases they evidently do. Even in the latter, however, we do not necessarily find that the pattern of disability corresponds strictly to a lesion of one or other hemisphere. The plasticity of the child's brain, and the fact that the brain damage was sustained before linguistic or visuo-motor skills had been acquired, renders the position very different from what is found in the brain-injured adult.

Perhaps I should say a word or two more about dyslexia, if only in view of the controversy which any mention of this term is liable to arouse. In the first place, it is important to bear in mind that, even in cases in which no definite evidence of brain damage can be elicited, the difficulties often resemble closely what is found in organic conditions (cf. Critchley, 1964). For example, these children have obvious difficulty in retaining serial order, a factor which, when taken together with poor rote learning, evidently contributes to the quite excessive difficulty in learning to spell (Doehring, 1968). In the second place, a family history of language backwardness is not uncommon, as also a familial history of left-handedness (Zangwill, 1960*a*). These factors, together with the tendency to reversals in reading and spelling, and sometimes to right-to-left ocular scanning (Blakemore & Zangwill, 1972), make one wonder whether a constitutional failure to establish cerebral dominance may not play some part in the condition. Whatever its ultimate nature, however, I would plead in the strongest terms that developmental disorders be approached within the general framework of paediatric neurology and genetics and that their existence, even when relatively mild, be taken seriously by psychologists, especially those whose special concern lies in education. There is no truth in the innuendo that dyslexia is a 'middle class disease', a mere name thought up by ambitious parents to disguise the stupidity or indolence of their offspring. Dyslexia may affect children, above all boys, of all classes and all levels of intelligence. It merits understanding and compassion, and above all research into its causes and remedial management at home and at school.

VII. Is there a take-over?

Is an approach to psychology such as I have outlined in the best interests of the subject? Do we really need to consider the brain in our attempts to explain human behaviour? These are questions which have long proved controversial. G. F. Stout reminded us in the first chapter of his *Manual of Psychology* that, had it not been for the anatomists, 'Conscious individuals might have thought, felt, willed and known about themselves without ever supposing that they had brains at all' (Stout, 1924, p. 17). In this, if only in this, Freud might well have agreed with him. Although everyone today knows that he has a brain, and few would deny that it has something to do with how he thinks, feels, wills and knows about himself, psychologists still disagree as to the relevance of physiology to their subject.

This rather central issue has recently been brought to the fore by Dr R. B. Joynson, a distinguished member of the Psychology Department at Nottingham University. In two papers published in our *Bulletin* (Joynson, 1970, 1972), followed by a short book (Joynson, 1974), this author has expressed his concern about a number of trends in contemporary psychology, not least the tendency to seek a coalescence between neurology and psychology, as advocated by the late Karl Lashley (1941) and echoed on occasion by myself. Although Joynson has no wish to deny that the relations between neurology and psychology should be close, he considers that it behoves those who are keen to foster it to provide a clear and positive indication of the role of psychology in any such alliance. 'In particular', he asks, 'does Zangwill have any workable conception of a psychological explanation as something in any way different from a physiological explanation?' Unless he does, Joynson concludes, the conception of a unified scientific discipline appears no more than a means of concealing the disappearance of psychology.

I am no philosopher and have no wish to get embroiled with the mind–body problem. But Joynson has offered a clear challenge and it is incumbent upon me to try to meet it. First, I should like to say what I understand by a physiological explanation: It is an explanation put forward by physiologists, whose discipline is traditionally concerned with the physical activities of the body. In so far as it has concerned itself with the nervous system, it has, in my view quite properly, resolutely set its face against explanation in terms of vital forces or psychical powers. In this connexion, one must bear in mind that the concept of a 'spinal cord soul' persisted right up to the end of the last century and that the notion of 'higher brain centres', which still finds a place in physiological texts, is an odd amalgam of physiological and psychological thinking (Zangwill, 1971). Given this background, it is hardly surprising that modern physiologists have adopted Cartesian dualism as their working philosophy and have been wary of dabbling in the realm of mind.

But today there are signs that the scene is changing. No longer do neurophysiologists confine their studies to the intimate biophysics of nervous transmission or to the electrical activities of single units. In studies of the central nervous system, more particularly those concerned with the functions of the hypothalamus, the limbic system and the cerebral cortex, the neurophysiologist is becoming increasingly involved in behavioural issues. And far from displacing the psychologist, he is increasingly welcomed as an essential partner. The psychologist brings to neurobiology not only the techniques of comparative psychology but also a body of thinking concerned with the nature of appetite, emotion, motivation and learning with which the physiologist is in general ill acquainted. Indeed there are already signs that theory in neurobiology is becoming a great deal more sophisticated than it once was and increasingly ready to embody what have traditionally been viewed as psychological modes of explanation.

I thus came to Joynson's central question. Do I have any workable conception of a psychological explanation as something in any way different from a physiological explanation? My answer can only be that I regard a psychological explanation as one which purports to account for the behaviour of an organism as a whole in its transactions with the environment. As I see it, the central nervous system is organized on a hierarchical basis and no clear-cut distinction can be drawn between those of its activities which are associated with consciousness and those which are not. Reflex adjustments merge into habitual reactions; habitual reactions into motor skills; motor skills into productive thinking. Psychology, moreover, unlike physiology, is intimately concerned with the issues of development. As J. B. Watson once put it: 'The physiologist *qua* physiologist knows nothing of the total situations in the daily life of an individual that shape his action and conduct. He may teach us all there is to know about the mechanism of stepping, but it is not his task to determine whether man walks before he crawls' (Watson, 1929, p. 20).

As I see it, then, a psychological explanation is one concerned with the development, organization and control of human behaviour. It should proceed – and here I believe I follow Fechner – in a manner consistent with our knowledge of the body and brain and loyal to the methods of the natural sciences. If so envisaged, I have no fear of that breakdown of modern psychology which Joynson so gloomily prognosticates. Let me end by citing the three closing words of Kenneth Craik's brilliant essay on *The Nature of Explanation* (1943), written in the darkest days of the last war when very much more than psychology seemed in imminent danger of dissolution. 'And so', wrote Craik, '*tentare*'.

REFERENCES

BARTLETT, F. C. (1932). *Remembering*. London: Cambridge University Press.

BARTLETT, F. C. (1957). *Thinking: An Experimental and Social Study*. London. Allen & Unwin.

BLAKEMORE, C. B., IVERSEN, S. D. & ZANGWILL, O. L. (1972). Brain functions. *Ann. Rev. Psychol.* **23**, 413–456.

BLAKEMORE, C. B. & ZANGWILL, O. L. (1972). Dyslexia: Reversal of eye movements during reading. *Neuropsychologia* **10**, 371–373.

BOGEN, J. E. & GAZZANIGA, M. S. (1965). Cerebral commissurotomy in man: Minor hemisphere dominance for certain spatial functions. *J. Neurosurg.* **23**, 394–399.

BORING, E. G. (1929). *A History of Experimental Psychology*. New York: Century.

BRENNER, N. W., GILLMAN, S., FARRELL, M. & ZANGWILL, O. L. (1967). Visuo-motor disability in schoolchildren. *Br. med. J.* **4**, 259–262.

COHEN, J. (1970). *Homo Psychologicus*. London: Allen & Unwin.

COLE, M. & ZANGWILL, O. L. (1963). *Déjà vu* in temporal lobe epilepsy. *J. Neurol. Neurosurg. Psychiat.* **26**, 37.

Craik, K. J. W. (1943). *The Nature of Explanation.* London: Cambridge University Press.

Critchley, M. (1964). *Developmental Dyslexia.* London: Heinemann.

Doehring, D. G. (1968). *Patterns of Impairment in Specific Reading Disability.* Bloomington: Indiana University Press.

Eccles, J. C. (1965). *The Brain and the Unity of Conscious Experience.* 19th Eddington Memorial Lecture. London: Cambridge University Press.

Eccles, J. C. (1973). *The Understanding of the Brain.* New York. McGraw-Hill.

Elvin, M. B. & Oldfield, R. C. (1951). Disabilities and progress of a dysphasic university student. *J. Neurol. Neurosurg. Psychiat.* **14**, 118–128.

Ettlinger, G., Jackson, C. V. & Zangwill, O. L. (1956). Cerebral dominance in sinistrals. *Brain* **79**, 569–588.

Ettlinger, G., Warrington, E. & Zangwill, O. L. (1956). A further study of visual-spatial agnosia. *Brain* **80**, 335–361.

Fechner, G. T. (1860). *Elemente der Psychophysik,* vol. 2. Leipzig: Breitkopf & Härtel.

Gardner, H. (1975). *The Shattered Mind: The Person after Brain Damage.* New York: Alfred A. Knopf.

Gazzaniga, M. S. & Sperry, R. W. (1967). Language after section of the cerebral commissures. *Brain* **90**, 131–148.

Hartmann, E. von (1931). *Philosophy of the Unconscious: Speculative Results According to the Inductive Method of Physical Science.* London: Kegan Paul, Trench & Trubner.

Hudson, L. (1966). *Contrary Imaginations.* London: Methuen.

Humphrey, M. E. & Zangwill, O. L. (1951). Cessation of dreaming after brain injury. *J. Neurol. Neurosurg. Psychiat.* **14**, 322–325.

Humphrey, M. E. & Zangwill, O. L. (1952). Dysphasia in left-handed patients with unilateral brain lesions. *J. Neurol. Neurosurg. Psychiat.* **15**, 184–193.

James, W. (1890). *Principles of Psychology,* vol. 2. London: Macmillan.

Joynson, R. B. (1970). The breakdown of modern psychology. *Bull. Br. psychol. Soc.* **23**, 261–269.

Joynson, R. B. (1972). The return of mind. *Bull. Br. psychol. Soc.* **25**, 1–10.

Joynson, R. B. (1974). *Psychology and Common Sense.* London: Routledge & Kegan Paul.

Köhler, W. (1925). *The Mentality of Apes.* Transl. by E. Winter. New York: Harcourt Brace.

Lashley, K. S. (1941). Coalescence of neurology and psychology. *Proc. Am. Philos. Soc.* **84**, 461–470.

Levy-Agresti, J. & Sperry, R. W. (1968). Differential perceptual capacities in major and minor hemispheres. *Proc. natn. Acad. Sci. U.S.A.* **61**, 1151.

Levy, J., Trevarthen, C. & Sperry, R. W. (1972). Perception of bilateral chimeric figures following hemispheric deconnection. *Brain* **95**, 61–78.

Luria, A. R. (1968). The mechanism of 'dynamic aphasia'. *Found. Language* **4**, 296–307.

McDougall, W. (1911). *Body and Mind.* London: Methuen.

McFie, J., Piercy, M. F. & Zangwill, O. L. (1950). Visual-spatial agnosia associated with lesions of the right cerebral hemisphere. *Brain* **73**, 167–190.

McFie, J. & Zangwill, O. L. (1960). Visuo-constructive disorders associated with lesions of the left parietal lobe. *Brain* **83**, 243–260.

Milner, B. (1962). Laterality effects in audition. In V. B. Mountcastle (ed.), *Interhemispheric Relations and Cerebral Dominance.* Baltimore: Johns Hopkins Press, pp. 177–195.

Newcombe, F. (1969). *Missile Wounds of the Brain. A Study of Psychological Deficits.* London: Oxford University Press.

Ornstein, R. (1972). *The Psychology of Consciousness.* San Francisco: W. H. Freeman.

Paterson, A. & Zangwill, O. L. (1944). Disorders of visual perception associated with lesions of the right cerebral hemisphere. *Brain* **67**, 331–358.

Piercy, M. F. & Smyth, V. O. G. (1962). Right hemisphere dominance for certain non-verbal intellectual skills. *Brain* **85**, 775–790.

Piéron, H. (1923). *Le Cerveau et la Pensée.* Transl. C. K. Ogden. *Thought and the Brain* 1927. London: Kegan Paul, Trench & Trubner.

Smith, A. (1966). Speech and other functions after left (dominant) hemispherectomy. *J. Neurol. Neurosurg. Psychiat.* **29**, 467–471.

Smith, A. (1972). Dominant and non-dominant hemispherectomy. In W. L. Smith (ed.),*Drugs, Development and Cerebral Function.* Springfield, Ill.: Thomas.

SPERRY, R. W., GAZZANIGA, M. S. & BOGEN, J. E. (1969). Interhemispheric relationships: The neocortical syndromes and hemisphere disconnection. In P. H. Vinken & G. W. Bruyn (eds), *Handbook of Clinical Neurology*, vol. 4. Amsterdam: North Holland.

STOUT, G. F. (1924). *A Manual of Psychology*, 3rd ed. London: University Tutorial Press.

WATSON, J. B. (1929). *Psychology from the Standpoint of a Behaviorist*, 3rd ed., revised. Philadelphia: Lippincott.

WEISENBERG, T. & MCBRIDE, K. E. (1935). *Aphasia: A Clinical and Psychological Study*. New York: The Commonwealth Fund.

ZANGWILL, O. L. (1960*a*). *Cerebral Dominance and its Relation to Psychological Function*. Henderson Trust Lecture No. 19. Edinburgh: Oliver & Boyd.

ZANGWILL, O. L. (1960*b*). Le problème d'apraxie idéatoire. *Rev. Neurol.* **102**, 596–603.

ZANGWILL, O. L. (1961). Asymmetry of cerebral hemisphere function. In H. Garland (ed.), *Scientific Aspects of Neurology*, pp. 51–62. Edinburgh: Livingstone.

ZANGWILL, O. L. (1964*a*). The current status of cerebral dominance. *Proc. Ass. Res. nerv. ment. Dis.* **42**, 103–118.

ZANGWILL, O. L. (1964*b*). Intelligence in aphasia. In A. V. S. de Reuck & M. O'Connor (eds), *Disorders of Language*, pp. 261–274. London: Churchill.

ZANGWILL, O. L. (1966). Psychological deficits associated with frontal lobe lesions. *Int. J. Neurol.* **2**, 395–402.

ZANGWILL, O. L. (1971). Diagram-makers old and new. *Totus Homo, Rome*, **3**, 53–58.

ZANGWILL, O. L. (1972). 'Remembering' re-visited. 3rd Sir Frederic Bartlett Lecture. *Q. Jl exp. Psychol.* **24**, 123–138.

ZANGWILL, O. L. (1974). Consciousness and the cerebral hemispheres. In S. J. Dimond & J. G. Beaumont (eds), *Hemisphere Function in the Human Brain*, pp. 264–278. London: Elek Science.

(*Manuscript received* 14 *August* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 315–325
Printed in Great Britain

THE EFFECTS OF STIMULUS ATTRIBUTES UPON LATENCY OF WORD RECOGNITION

BY JOHN T. E. RICHARDSON*
Neuropsychology Unit, Department of Neurology, The Churchill Hospital, Oxford

This experiment investigated the properties affecting the time taken to read individual words and to discriminate between words and non-words. The number of letters in a word was found to affect the time taken to read it aloud, but not the time taken to discriminate it from non-words. The frequency of a word affected the time for discrimination, but not that for reading. Non-words which obeyed the rules of English orthography and phonology were more difficult to discriminate from words than those which violated those rules. Imageability, concreteness, and the number of syllables in a word were found not to affect performance, nor were derived nouns more difficult to process than simple nouns. It is suggested that reading aloud employs grapheme–phoneme translation based upon a letter-by-letter analysis of the stimulus; that discriminating words from non-words obeying the rules of English orthography and phonology employs a search of the lexicon based upon a holistic analysis of the stimulus; and that discriminating words from non-words violating those rules employs a direct test of the regularity of the stimulus based upon the combinatory rules of English orthography.

This study investigated the effects of stimulus attributes in word recognition, using two reaction-time situations: in the *naming* task, the subject has to read a single word aloud as quickly as possible; in the *classification* task, the subject has to decide whether or not a set of letters constitutes a word. Although previous studies have considered the particular problems investigated here, they do not allow one to come to any firm conclusions. The principal reason for this is that experimenters have consistently failed to employ the correct statistical techniques for generalizing over both subjects and stimulus material (Clark, 1973). Their conclusions hold only for the particular sets of stimuli employed, and there exists, therefore, no reliable general evidence concerning the effects of stimulus attributes in word-recognition tasks. The experiment reported here attempted to remedy this situation.

The first question with which this study was concerned is whether semantic factors are involved in the reading and recognition of individual words. The particular attributes of interest were imageability and concreteness, since it is a consequence of Paivio's (1971, chapter 4) two-process theory of imagery and language that these attributes should not affect the immediate perceptual recognition of a word. Experiments which have employed tachistoscopic presentation to investigate this issue have produced inconsistent results (Riegel & Riegel, 1961; Winnick & Kressel, 1965; Paivio & O'Neill, 1970). Using a classification task, Rubenstein, Garfield & Millikan (1970) found no difference in the reaction times to concrete and abstract words (cf. Clark, 1973).

The second question was whether possible non-words (that is, groups of letters which satisfy the rules of English orthography and phonology, yet which do not belong to the English language, e.g. *stul*) are processed in a different manner from impossible non-words (groups of letters which violate those rules). Gibson, Pick,

---

* Now at Department of Psychology, Brunel University, Uxbridge, Middx. UB8 3PH.

Osser & Hammond (1962) and Gibson, Bishop, Schiff & Smith (1964) suggested that possible non-words were correctly reported at shorter tachistoscopic durations than impossible non-words. In classification experiments, Rubenstein, Lewis & Rubenstein (1971) and Snodgrass & Jarvella (1972) concluded that impossible non-words were rejected faster than possible non-words.

The third question was whether words with suffixes are processed in some special manner. This is of theoretical interest, since words such as *hardness* may be regarded as lexically derived from their stems, and the appropriate explanation of this derivation is a topic of some controversy in linguistics (see, for example, Chomsky, 1970). Snodgrass & Jarvella (1972) investigated the effect of adding either a prefix or a suffix to a basic stem, yielding words such as *preschool* and *hardness*, possible non-words such as *prestul* and *diltness*, and impossible non-words such as *prebkdut* and *btulness*. Unfortunately, in their experiment the presence of an affix was completely confounded with an increase in the number of letters and syllables in the stimulus, and a change in its frequency (if it was a word) or in its degree of approximation to English (if it was a non-word).

The fourth problem to be investigated is the separation and interpretation of the effects of word frequency and of word length. The latter factor may be studied either from the point of view of orthography (i.e. the number of letters) or from the point of view of phonology (e.g. the number of syllables). Many investigators have attempted to separate these variables in a variety of situations (e.g. McGinnies, Comer & Lacey, 1952; Postman & Adis-Castro, 1957; Oldfield & Wingfield, 1965; Eriksen, Pollack & Montague, 1970; Berry, 1971; Klapp, 1971; Klapp & Bischoff, 1972; Carroll & White, 1973; Forster & Chambers, 1973; Kalpp, Anderson & Berrian, 1973; Spoehr & Smith, 1973). However, their conclusions are inconsistent with one another, and the analyses of results are subject to the criticisms presented by Clark (1973).

METHOD

Subjects

Thirty undergraduates at the University of Sussex were allocated at random to three equal groups; each group participated in one of the three experimental conditions.

Material

Eight lists of ten nouns were produced, which orthogonally varied high vs. low imageability, high vs. low concreteness, and simple vs. derived nouns. Imageability (I) and concreteness (C) were taken from the norms published by Paivio, Yuille & Madigan (1968), but the selection of material was validated by *post hoc* rating studies using English students, which suggested that the basic design was not affected by the difference between English and Canadian raters. A derived noun was defined as a noun consisting of an explicit noun-forming suffix, together with a stem which was identical to, or was a minimally modified form of, an English word in common use. Any noun which did not satisfy this criterion was regarded as simple. High imageability was defined by I being greater than 3·68, low imageability by I being less than 3·68. High concreteness was defined by C being greater than 3·0, low concreteness by C being less than 3·0. Words in corresponding lists were matched on the basis of I and C, and also on the basis of frequency (Thorndike & Lorge, 1944); the latter varied between 1 and AA, and lists had median values between 13 and 15. Matching was approximate in the case of individual items and more exact in the case of list means. To control for the length of a word in terms of its number of syllables, three-syllable words were used as far as possible, but seven words of two syllables and

Table 1. *Examples of experimental material*

		Words	Possible non-words	Impossible non-words
Simple nouns				
High *I*	High *C*	DISASTER	SEDISART	SDEIATRS
	Low *C*	DELIRIUM	IMURILED	RLIMEIUD
Low *I*	High *C*	CHARLATAN	CANARLATH	RCLAANHTA
	Low *C*	ECONOMY	MYONOCE	NCOOEYM
Derived nouns				
High *I*	High *C*	POLLUTION	OLLUPTION	LPUTLOION
	Low *C*	HAPPINESS	PAPHINESS	NHIAPSESP
Low *I*	High *C*	SOCIALIST	ALICOSIST	CSOAITSIL
	Low *C*	DISTINCTION	STINDICTION	DOTINCIINST

five words of four syllables also had to be included, as sufficient material did not exist in the published norms of *I* and *C*. The number of letters in a word varied from 5 to 11, but the mean number of letters did not vary systematically between the eight lists.

Two non-word anagrams were constructed from each word. One anagram was constructed so as to be a possible non-word; in this case, the anagram of a derived noun retained the noun-forming suffix. The other anagram was constructed so as to be an impossible non-word. Examples of the various types of material are shown in Table 1.

Procedure

For the naming task, Condition A, the stimuli were the 80 words mentioned above, and 40 other similar words as practice items. The critical items were presented in five blocks, each containing four practice items and two words from each of the eight lists. The remaining 20 practice items formed a sixth block. The stimuli were typed in upper-case letters on IBM cards, which were presented in six stacks mounted horizontally behind a shutter. The subject was instructed to say 'Now!' into a hand-held microphone, whereupon a Venner Mark I electromagnetic clock was triggered, and the shutter rose, revealing the stimulus. The subject pronounced the word, and his response stopped the clock and returned the shutter. The experimenter removed the top card and recorded the latency of the response. (This latency was the duration between the termination of the subject's 'Now!' and the initiation of his overt response.) Every subject received the block of practice items first (though they were not identified as such), and the order of the other five blocks to the ten subjects was determined by a Latin square. The order of presentation of the items within each block was constant.

For the first classification task, Condition B, the stimuli were the 80 critical words, the possible non-word anagrams of those words, 30 similar practice words, and the similarly constructed anagrams of those words. The experimental words were combined into ten sets on the basis of equal rank frequency within each list; that is, the most frequent items in each list was put into one set, the second most frequent into another, and so on. Ten sets of anagrams were constructed, such that each set consisted of the anagrams of the items in one set of words. The sets of words and non-words were paired at random, with the constraint that corresponding words and non-words were not paired together. Each of the ten sets of 16 stimuli thus formed was then randomized, with the constraint that no successive set of four items was either all words or all non-words, and was preceded by two practice words and two practice non-words, obeying the same constraints. A practice set was constructed of ten words and ten non-words, similarly randomized. The 11 blocks of 20 items were typed in upper-case letters on IBM cards, and the procedure was identical to that of Condition A, except that the subjects were to respond 'Yes' to words and 'No' to non-words. Every subject received the block of practice items first (though they were not identified as such), and the order of the other ten blocks to the ten subjects was determined by a Latin square. The order of presentation of the items within each block was constant.

The second classification task, Condition C, was identical to Condition B, except that the impossible non-word anagrams were used, and, if the apparatus failed to register the subject's response to a given stimulus, that stimulus was presented for a second time at the end of its block.

Table 2. *Harmonic mean latencies* (*msec*)

		Words					
		High *I*		Low *I*			
		High *C*	Low *C*	High *C*	Low *C*	Overall	Non-words
Condition A	Simple	596	554	576	560	571	
	Derived	554	565	561	597	569	
	Overall	574	559	569	578	570	
Condition B	Simple	662	641	698	662	665	713
	Derived	640	667	675	669	663	739
	Overall	651	654	686	666	664	726
Condition C	Simple	595	585	590	587	589	
	Derived	563	573	596	569	575	
	Overall	578	579	593	578	582	598

Results

The data for analysis were the latencies from all those instances where a subject gave a correct response which triggered the apparatus. In Condition A, there were 67 missing observations (8·4 per cent); these were all caused by the apparatus failing to register the subject's response, i.e. there were no cases where an incorrect response was registered by the apparatus. In Condition B, there were 145 missing observations (9·1 per cent). These included both failures of the apparatus and errors on the part of the subject, i.e. incorrect responses which triggered the apparatus; the error rates for words and for non-words were 5·8 and 6·8 per cent, respectively, and the failure rates were 2·9 and 2·8 per cent, respectively. In Condition C, there were no failures of the apparatus contributing to the 44 missing observations (2·8 per cent), and the error rates for words and for non-words were 3·6 and 1·9 per cent, respectively. Although the numbers of errors was too small for any detailed analysis, a Mann–Whitney U test on each subject's percentage of errors (out of his total number of errors and correct responses) showed that, for this material, subjects in Condition B perform worse than subjects in Condition C ($U = 13$, $n_1 = n_2 = 10$, $P < 0{\cdot}01$, two-tailed test). The raw data in this experiment were subjected to the reciprocal transformation, and the missing observations were replaced by a procedure described by Winer (1962, p. 281), whose chief assumption is that there is no interaction between subjects and stimuli. Harmonic mean latencies are shown in Table 2.

A separate analysis of variance was carried out on the transformed data from each condition. In all three cases, the analysis of the latencies on the 80 critical words contained the factors of imageability, concreteness, and lexical complexity (i.e. simple vs. derived nouns). The appropriate procedures were used for generalizing over both subjects and stimuli (cf. Clark, 1973). For the analysis of Condition A, no term was significant. For Condition B, in the case of the non-words, two factors were used: whether the item was the anagram of a simple or a derived noun, and the rank frequency of the words in the context of which the non-word was presented. (Only the linear trend over rank frequency was considered to be of theoretical interest.) Two factors were significant: subjects took longer to reject non-words than

Table 3. *Partial correlation coefficients between the reciprocal of the harmonic mean response latency on each word in each condition, and each of five stimulus attributes*

($|r| > 0{\cdot}25$, $P < 0{\cdot}05$)

Stimulus attributes	Conditions A	B	C
I	−0·08	0·21	0·08
C	−0·11	−0·23	−0·17
m	0·13	0·10	−0·01
F	0·19	0·54	0·27
L	−0·36	−0·04	0·12

to accept words (quasi $F = 24{\cdot}03$, d.f. $= 1, 24$, $P < 0{\cdot}001$); and the linear trend over rank frequency in the case of non-words was also significant (quasi $F = 5{\cdot}01$, d.f. $= 1, 32$, $P < 0{\cdot}05$), such that non-words in the context of higher-frequency words were rejected faster than those in the context of lower-frequency words (717 msec for the five blocks containing higher-frequency words, and 735 msec for the remainder). For Condition C, in the case of the non-words, only the rank frequency of the words in the context of which the non-word was presented was considered. No comparison was significant. In none of the three conditions did any of the three properties of the words yield a significant effect. A Mann–Whitney U test on each subject's harmonic mean response latency showed that, for this material, subjects in Condition C accept words faster than subjects in Condition B ($U = 12$, $n_1 = n_2 = 10$, $P < 0{\cdot}01$, two-tailed test).

The effect of word length in terms of number of syllables was investigated by means of analyses of covariance which used a single fixed factor of number of syllables and five covariates. These were imageability (I), concreteness (C), meaningfulness (m, taken from Paivio *et al.* 1968), frequency (F, taken from Thorndike & Lorge, 1944; 'AA' words were scored as 100, 'A' words as 50), and length in terms of the number of letters in a word (L). In Condition A, the harmonic mean response latencies for two-, three-, and four-syllable words after adjustment for the effects of the covariates were 597 msec, 560 msec, and 576 msec. In Condition B, the corresponding means were 666 msec, 662 msec, and 667 msec, and in Condition C, they were 588 msec, 585 msec, and 587 msec. In none of the conditions was the effect of number of syllables significant ($F < 1$ in all cases). Consequently, correlation analyses were carried out, ignoring the classification by number of syllables, on the five covariates and the reciprocal of the harmonic mean latency on each word over the ten subjects in each condition. Table 3 shows the partial correlation coefficients with each stimulus attribute, controlling for the effects of the other four attributes. When such effects are statistically controlled, only length in terms of number of letters affects the response latency in Condition A, and only frequency affects the response latency in Conditions B and C.

The non-words in Condition B were analysed by an analysis of covariance which used a criterion of number of syllables and a covariate of number of letters. The adjusted harmonic means were: two syllables, 733 msec ($n = 10$); three syllables, 722 msec ($n = 53$); and four syllables, 733 msec ($n = 17$). The correlation between number of letters and the reciprocal of the harmonic mean latency was $-0{\cdot}005$. Neither the criterion nor the covariate yielded significant effects ($F < 1$ in both cases). For the non-words in Condition C, the correlation between number of letters and the reciprocal of the harmonic mean latency was $-0{\cdot}016$, which is not significant ($F < 1$).

DISCUSSION

The first problem investigated here was whether the time taken to read or recognize a word depends upon its imageability or concreteness. The tasks employed in this experiment involved the perception, recognition, classification, and articulation of individual words, and yet no evidence has been found to suggest that either imageability or concreteness is involved in any of these processes. This is consistent with the theoretical analysis of imagery and language given by Paivio (1971), though of course the general theoretical position that a word's meaning does not affect the ease with which it may be recognized has a much older history.

The second problem was whether possible non-words are treated in a different manner from impossible non-words. In the present experiment, distinguishing words and possible non-words was found to be a more difficult task than distinguishing words and impossible non-words, both in terms of the subjects' accuracy and in terms of their latency of responding. These results do not however entail that subjects employ the rules of English orthography and phonology in order to reject impossible non-words, or that rejecting possible non-words and rejecting impossible non-words involve different processes in any way. It is possible that in both tasks subjects make some decision concerning the familiarity of the combination of letters which constitutes a stimulus, a decision which is easier in the case of impossible non-words and common words, but harder in the case of possible non-words and rare words. Nevertheless, on this account one ought to expect a difference between possible non-words with a noun-forming suffix and those without such a suffix, for such suffixes occur frequently in written language. If the subjects were making some decision about the familiarity of the stimulus, therefore, it is surprising that no difference was found between these two classes of stimulus. One might further expect a difference between the simple nouns and the derived nouns, for the latter are composed of more frequent combinations of letters, even though the frequency of the whole word was controlled. Yet no difference was found in either Condition B or Condition C. It is therefore easier to account for the difference in performance between the possible non-words and the impossible non-words in terms of their legality, rather than in terms of their familiarity. This will be considered in more detail presently.

The above discussion relates to the third problem, which was whether words with suffixes are processed in some special manner. No evidence was found in this experiment to suggest that words with noun-forming suffixes were processed in a different manner from words without such suffixes, nor that possible non-words with noun-forming suffixes were processed in a different manner from possible non-words without such suffixes. The importance of the latter finding has already been considered. The fact that there was no effect in the case of words counts against the view that derived nouns are linguistically more complex than simple nouns. One might try to explain it by supposing that the simple and derived forms of a word (e.g. *hard* and *hardness*) have a common lexical representation, which is neutral between the grammatical categories of the two forms. This claim can be supported by arguments from linguistics (Chomsky, 1970). However, on this account, the total frequency of occurrence of the base lexical item would be greater than that of either of its surface forms. Thus, if an experiment produces an effect of frequency of occurrence, it should also produce a difference between simple and derived nouns even when the words themselves are matched on that attribute. The fact that there is no difference between performance on simple and derived nouns even in those conditions where there is an effect of frequency counts against the idea that there is any psychologically relevant difference between the two classes of noun.

The fourth problem was the separation of the effects of three attributes of the stimulus material. In this experiment, the frequency of a word was found to be important in the classification task, but not in the naming task; the number of letters in a word was important in the naming task, but not in the classification task; and the number of syllables in a word was not important in either task. These results contradict the claim made by Forster & Chambers (1973) that 'naming time is generally influenced by the same variables that affect lexical decision time'. Indeed, the dissociation between the experimental paradigms and the stimulus attributes indicates that the properties of word frequency and word length relate to different psychological processes. The results also suggest that the latter property only relates to performance insofar as it is measured by the number of letters in a word; the number of syllables does not predict performance in any of the situations studied here. (Of course, this does not rule out the possibility that the number of syllables may have an effect in tasks using other vocabularies, e.g. two-digit numbers.) The effects of length and frequency and the structures of the experimental tasks may be discussed in more detail.

One process which is involved in the naming task but not in the classification task is the programming of the speech mechanism to produce sounds appropriate to the particular stimulus. (The latencies recorded do not include the time for the overt articulation of a given word.) One might therefore conclude that longer words take longer to name because they involve a more complex articulatory programme. This is the approach adopted by Klapp *et al.* (1973). However, in order to explain the present results in this manner, one would have to make the counterintuitive claim that the number of letters in a word is a better measure of its phonological complexity than the number of syllables. This approach would also predict that word length would affect performance whenever a word had to be articulated. Object-naming is such a task, and yet the number of letters in an object's name has never been found to correlate with the naming latency (Oldfield & Wingfield, 1965; Carroll & White, 1973; cf. Klapp *et al.* 1973). The number of letters is more obviously an index of orthographical complexity, rather than phonological complexity, and a more satisfactory explanation of the present results would be to suggest different strategies of perceptual analysis. That is, subjects appear to adopt a holistic analysis of the stimulus in the classification task, but they adopt a letter-by-letter approach in the naming task.

The finding that frequency affects the time taken to classify a word, but not the time taken to read it aloud, suggests that some further processing is involved in classifying beyond whatever is sufficient for naming. If one considers just the task employed in Condition B, the only conclusive way to classify a word is to find a representation in the lexicon, to find a meaning for it. One might therefore take the results of this condition to suggest that the lexical address of a common word may be accessed faster than that of a rare word. Further, the only conclusive way to reject a possible non-word is to search the lexicon in an exhaustive fashion before responding, and this would explain the longer reaction times for non-words as compared to words. The effect of frequency may be handled in at least two ways. First, one might suppose that lexical search is directed by frequency of occurrence, so that the addresses of common words are tested before those of rare words. Second, a preliminary analysis might be employed to extract information from the stimulus which would ascribe it to a particular frequency class, and the search could then be confined to the section of the lexicon relevant to that class (Oldfield, 1966). The former account can explain the effect of frequency in terms of the serial nature of lexical search; the latter account can explain it by assuming that a finer preliminary analysis is possible in the case of common words, and that this results in a smaller set of lexical entries to be tested. Berry (1971) found that words could be named more quickly in lists which were homogeneous, rather than heterogeneous, with respect to frequency of occurrence. He considered that this was evidence for a two-stage model of lexical search, but the results of his experiment are not supported by the present study, which failed to find an effect of frequency in naming. Nevertheless, the present finding, that non-words are rejected faster in the context of words of higher frequency, suggests that fewer lexical items have to be examined in order to make such a decision, and thus supports the notion that the classification of common words utilizes a smaller search set than that of rare words. (An explanation of this finding in terms of a judgement of familiarity would of course have to handle the difficulties raised in the second paragraph of this section.)

However, the results of the classification task using possible non-words differ in several respects from those of the task using impossible non-words. The average latency of accepting a word in Condition B is 80 msec longer than that in Condition C, and this suggests that the two conditions may involve different processes. Although frequency of occurrence bears a significant relationship to performance in both tasks, this correlation is significantly smaller in Condition C than in Condition B ($z = 1{\cdot}98$, $P < 0{\cdot}05$, two-tailed test). In terms of the above explanation of the frequency effect in Condition B, this suggests that at least some words are accepted without lexical search when they are presented in the context of impossible non-words. Indeed, if, as suggested earlier, subjects can employ information about the legality of the stimulus, lexical search is quite unnecessary. That impossible non-words are not rejected on the basis of an exhaustive search of the lexicon is supported by the fact that the time to make such a decision is no longer than the time to accept a word. This cannot be explained in terms of different response latencies for the words 'yes' and 'no': in a control task where stimuli were presented 2 sec before a signal to respond, Forster & Chambers (1973) found that the latency to say 'yes' was 305 msec, and the latency to say 'no' was 310 msec. In Condition C, therefore, words and non-words seem to require the same amount of processing time. Further, the latency for rejecting a non-word is not affected by the frequency of the words in the context of which it is presented (quasi $F < 1$); in terms of the above discussion, this suggests that non-words are not rejected in condition C by an exhaustive search of a particular frequency class of the lexicon. Of course, if non-words are not rejected by means of lexical search, it would be redundant to carry out such a search in order to accept words. Thus, several findings rule against the suggestion that the decisions made in Condition C are based upon lexical search, and one might conclude instead that subjects test the stimuli for orthographic or phonological regularity. The effect of frequency observed in this condition could be explained in terms of the complexity of the parsing needed to establish this regularity; that is, one might assume that rare words are more complex than common ones, either in terms of the combinatory rules defining possible combinations of letters, or in terms of the rules for translation from an orthographic to a phonological representation. Non-words can be rejected at any stage of the parsing process, and so it would not be expected that their classification time would exceed that for words.

Conclusion

The naming task employed in Condition A necessarily involves a translation from the orthographic information which is presented to the subject to the phonological information which is required of him. This was the only condition in which the length of a word showed a significant relationship to performance, and it was concluded above that it involves a perceptual analysis which proceeds letter-by-letter. One might therefore conclude that this kind of analysis is required when the task demands grapheme–phoneme translation. On the other hand, the classification tasks showed no effect of the length of a word, and it was suggested that they involve a holistic perceptual analysis. In distinguishing between words and possible non-words, subjects have to employ lexical search in order to arrive at a conclusive decision, and this search seems to involve the initial selection of a search set, based upon an estimate of the frequency of occurrence of the stimulus. Rubenstein *et al.* (1971) have suggested that lexical search employs a phonological recoding of the stimulus, but Clark (1973) has shown that the evidence which they present for this hypothesis is statistically unreliable. The results of the naming task employed here suggest that this would require a letter-by-letter analysis of the stimulus, and hence would produce an effect of the length of a word in the subjects' performance. (By the same token, the frequency judgement on which the search is held here to be based must also employ the orthographic representation of the stimulus.) Moreover, clinical studies suggest that lexical information may be available when phonological recoding is difficult or impossible (Marshall & Newcombe, 1973), and, in addition, several linguists have suggested that the rules of English orthography include devices to facilitate the direct lexical interpretation of written language at the expense of a more complex relation between the orthographic and phonological representations of a word or sentence (Dale, 1972, pp. 186–194). Thus, the lexical interpretation of orthographic stimuli seems to be carried out without phonological recoding. However, when subjects have to distinguish between words and impossible non-words, they do not appear to employ lexical search at all. Rather, they are able to test the regularity of a stimulus directly. Once again, if translation from orthography to phonology would require a letter-by-letter analysis of the stimulus, and hence would produce an effect of word length on performance, then such a process does not appear to be involved in the test of regularity; this employs a parsing based upon the orthographic combinatory rules.

This research was supported by the Science Research Council, and by Medical Research Council grant G973/144/C. I am grateful to Professor A. D. Baddeley, Dr M. Coltheart, Mr A. J. Marcel, and Dr Freda Newcombe for their comments on earlier drafts of this paper. The apparatus was constructed by Dr A. J. Wilkins.

REFERENCES

Berry, C. (1971). Advanced frequency information and verbal response times. *Psychon. Sci.* **23**, 151–152.

Carroll, J. B. & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quart. Jl exp. Psychol.* **25**, 85–95.

Chomsky, N. (1970). Remarks on nominalization. In R. A. Jacobs & P. S. Rosenbaum (eds), *Readings in English Transformational Grammar.* Waltham, Mass.: Ginn.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *J. verb. Learn. verb. Behav.* **12**, 335–359.

Dale, P. S. (1972). *Language Development: Structure and Function.* Hinsdale, Ill.: Dryden Press.

Eriksen, C. W., Pollack, M. D. & Montague, W. E. (1970). Implicit speech: Mechanism in perceptual encoding? *J. exp. Psychol.* **84**, 502–507.

Forster, K. I. & Chambers, S. M. (1973). Lexical access and naming time. *J. verb. Learn. verb. Behav.* **12**, 627–635.

Gibson, E. J., Bishop, C. H., Schiff, W. & Smith, J. (1964). Comparison of meaningfulness and pronunciability as grouping principles in the perception and retention of verbal material. *J. exp. Psychol.* **67**, 173–182.

Gibson, E. J., Pick, A. D., Osser, H. & Hammond, M. (1962). The role of grapheme–phoneme correspondence in the perception of words. *Am. J. Psychol.* **75**, 554–570.

Klapp, S. T. (1971). Implicit speech inferred from response latencies in same-different decisions. *J. exp. Psychol.* **91**, 262–267.

Klapp, S. T., Anderson, W. G. & Berrian, R. W. (1973). Implicit speech in reading, reconsidered. *J. exp. Psychol.* **100**, 368–374.

Klapp, S. T. & Bischoff, D. M. (1972). Does implicit speech in same-different decisions extend to nonsense forms? *Percept. Psychophys.* **11**, 363–364.

Marshall, J. C. & Newcombe, F. (1973). Patterns of paralexia. *J. psycholing. Res.* **2**, 175–199.

McGinnies, E., Comer, P. B. & Lacey, O. L. (1952). Visual-recognition thresholds as a function of word length and word frequency. *J. exp. Psychol.* **44**, 65–69.

Oldfield, R. C. (1966). Things, words and the brain. *Q. Jl. exp. Psychol.* **18**, 340–353.

Oldfield, R. C. & Wingfield, A. (1965). Response latencies in naming objects. *Q. Jl exp. Psychol.* **17**, 273–281.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

Paivio, A. & O'Neill, B. J. (1970). Visual recognition thresholds and dimensions of word meaning. *Percept. Psychophys.* **8**, 273–275.

Paivio, A., Yuille, J. C. & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. exp. Psychol.* Monogr. Suppl. **76**, 1–25.

Postman, L. & Adis-Castro, G. (1957). Psychophysical methods in the study of word recognition. *Science* **125**, 193–194.

Riegel, K. F. & Riegel, R. M. (1961). Prediction of word-recognition thresholds on the basis of stimulus-parameters. *Language & Speech* **4**, 157–170.

Rubenstein, H., Garfield, L. & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *J. verb. Learn. verb. Behav.* **9**, 487–492.

Rubenstein, H., Lewis, S. S. & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *J. verb. Learn. verb. Behav.* **10**, 645–657.

SNODGRASS, J. G. & JARVELLA, R. J. (1972). Some linguistic determinants of word classification times. *Psychon. Sci.* **27**, 220–222.

SPOEHR, K. T. & SMITH, E. E. (1973). The role of syllables in perceptual processing. *Cog. Psychol.* **5**, 71–89.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words*, New York: Bureau of Publications, Teachers' College.

WINER, B. J. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

WINNICK, W. A. & KRESSEL, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning, and immediate recall as a function of abstractness–concreteness and word frequency. *J. exp. Psychol* **70**, 163–168.

(*Manuscript received* 4 *November* 1974; *revised manuscript received* 9 *April* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 327–338
Printed in Great Britain

SEMANTIC INTERFERENCE FROM UNATTENDED PRINTED WORDS

GEOFFREY UNDERWOOD
University of Nottingham

Is attention necessary for reading? Whilst subjects named simple line-drawings which were exposed briefly in these experiments, irrelevant words were also displayed for short intervals. The naming latencies to the drawings varied as a function of attentional selectivity and as a function of the semantic association between picture and word. With focused attention only related words caused interference, but with a divided attention strategy unrelated words also retarded the naming latencies to the drawings. Words may be recognized whilst attention is diverted away from them, but if attention is highly selective then only semantically related items will affect performance. If the extraction of individual meanings of words is a preattentive process, then the selective admittance of information to awareness, in reading and elsewhere, may be for purposes of schematic assimilation.

One of the more engaging problems which experimental psychologists have encountered during the past 20 years has been that of the role of attention in behaviour. Attention may be necessary, for example, in the selection of drive-relevant information from the environment and for the organization of sensations into perceptual hypotheses, and this general position may be described as the early-selection theory of attention as propounded in a number of variations by Broadbent (1958), Treisman (1960) and Neisser (1967). By this model selectivity is governed primarily by the physical attributes of messages. Selection is said to be made prior to semantic analysis, and so the semantic attributes of unattended messages generally have no play in the determination of organized behaviour, unless there is a contextual relationship between attended and unattended words. The opposing view considers that perception is a preattentive process, that attention is required for the organization and execution of the response. Here selection takes place with the aid of both semantic and physical information about the signal sources, with the most pertinent signal gaining further processing to influence the response. This position may be described as the late-selection model, as advocated by Deutsch & Deutsch (1963) and Norman (1968).

Early experiments designed to test between these two main classes of theories of attention produced equivocal results, partly because each of the models assumes only one locus of selectivity (whereas an information-processing analysis allows for many – see Erdelyi, 1974), and also because of a basic similarity between statements of the models in that a process of admittance of information to awareness could be considered as 'perception' by the early-selection models and as 'response' by the late-selection models. Further faults with the initial experiments were methodological in nature. Characteristically these experiments tested for the perception or memory of an unattended message by first engaging the subject in a highly demanding task for which attention was thought to be essential (e.g. shadowing) and then asking the subject to make occasional responses to particular items presented in a second (unattended) message. The characteristically low rates of response to unattended words (e.g. Treisman & Geffen, 1967) may have been due to perceptual attenuation, but they could have been due to a response to the attended message precluding any response to any other message however well that other message may have been recognized, and however well integrated into the listener's schemata. Attention to one message may prevent any response to a second message being offered volitionally, and requesting subjects to respond to unattended messages may be unrealistic. Furthermore this experimental paradigm places the subject in a double-bind situation: at one time the experimenter tells the subject not to attend to a message, and then he tells the subject to respond to certain aspects of that same message. We cannot be sure that any successful responses to 'unattended' words do not follow a shift of attention and directed processing. The most frequently employed technique in attention research has been shadowing (the subject repeats back each word of the message as soon as he hears it), because it undoubtedly forces the unpractised subject to listen to the required message. The shadowing studies may provide a demonstration of the ability of subjects to report their experiences whilst attending closely to one aspect of their environment but it seems likely that the shadowing technique itself is responsible for producing a particular distribution of attention which is not necessarily typical of volitional activity. When subjects produce no continuous overt response during the presentation of dichotic messages, but still give evidence of selective listening, then they are able to respond to more of the secondary message than when required to shadow (Underwood & Moray, 1971; Underwood, 1972, 1973). Shadowing is a particularly difficult task for the unpractised subject, and the processing capacity necessary for its organization and execution may reduce the capacity which might otherwise be available for analysis of the secondary message. By this model the experience of attention may correspond to the allocation of the bulk of the available processing capacity to the 'attended' message (Underwood, 1976), and the necessity of performing the shadowing task will consume capacity and thereby affect adversely the processing of unattended messages.

More recent work has been avoiding these methodological faults, if not the theoretical problems, by testing subjects for the effects of unattended messages. In these experiments subjects are never tested directly for their reception of the unattended message, but reception is inferred from its effect upon the response to the attended message. This method allows a different perspective upon the description of the attention process, for attention is now seen as a process whereby irrelevant information may be excluded from the organization of the drive-related response. For example, Lewis (1970) found that the shadowing latency was dependent upon the semantic content of the unattended message in a dichotic listening task where no response to the unattended message was required. Similar experiments have been reported by Corteen & Wood (1972), by MacKay (1973) and by Smith & Groen (1974), demonstrating that the semantic content of unattended words may influence ongoing behaviour. These experiments go some way to supporting the late-selection theories of attention in that unattended words may contact their representations in memory, or excite their logogens (Morton, 1969) and hence affect other activities, without necessarily being available to awareness. Attention may be necessary for the organization of the available sensations into the singular perception upon which the response is dependent. Unattended words may excite their logogens and be available for assimilation into the individuals' schemata as a process of 'secondary perception' (Anderson, 1975), but attention may be required for the admittance of these words to consciousness where the assimilation process is directed. This is a similar view to the role of attention envisaged by Wundt (1897), in which lexical analysis may be pre-attentive but in which attention is required for the perception and elaboration of relations between words.

The two experiments to be described here follow the recent trend in the study of the attention process in that the unattended message was never tested, and so subjects never had cause to attend to it. The experiments also avoid the use of the shadowing procedure to control the direction of attention. The central question addressed by these studies concerns the effect of unattended printed words upon a picture-naming task to determine whether print can be read even when it is not being attended to. Whilst subjects attended to a briefly presented line-drawing, and named it, they were also shown a single word. If this unattended word affects the latency of naming the picture then we may conclude that attention is unnecessary for the extraction of meaning from print. There is by now a considerable body of evidence to suggest that a task-irrelevant word will affect the processing of an attended word, but little work has been done directly on the manipulation of attentional strategies during such processing. Jacobson (1973) found that a word which was backwards-masked by an associate was read more easily than a word masked by an unrelated word, and suggested that the target logogen was partially excited by the masking stimulus. Bradshaw (1974) has reported that the interpretation of an ambiguous word is likely to be affected by a disambiguating word presented at the same time, whether that word can be reported or not. In that experiment subjects were aware of the location of the target word, but not the location of the disambiguating word. Previous tasks completed by the subject can also affect the processing of print, and Conrad (1974) has demonstrated an effect of task-irrelevant words when subjects were unaware of the interpretations of ambiguous words. A sentence was first presented in which the final word was ambiguous. Following this oral presentation subjects saw a word printed in colour, and the task was to name the colour of ink used (i.e. a Stroop task). The time required to name the colour of ink was greater if the coloured word was semantically related to the ambiguous word at the end of the previous sentence than if it was unrelated, and moreover, a relationship slowed the response to the colour whichever way the ambiguous word had been interpreted by the sentence. Even when the second meaning of the word had not been indicated by the sentence a relationship between that meaning and the coloured word affected performance. Both meanings of an ambiguous word may be activated upon presentation. Posner & Snyder (1975), in reviewing such evidence of the perception of meaning, concluded that word recognition is an automatic process, and that attention is used primarily to reduce interference from unwanted sources. Attention is considered to facilitate processing, but is not essential for a message to be recognized. When an unattended message is presented and is effective the subject may well be unaware of the message or its effects, as demonstrated by the Conrad (1974) and Bradshaw (1974) experiments.

In the first experiment to be described here subjects were certain as to where in the visual field the line-drawing would appear, and in the second experiment they were uncertain. Optimal strategies may then have been to employ focused attention and divided attention respectively (Treisman, 1969). By the Posner & Snyder (1975) analysis we might expect that words related to the pictures would cause more interference than unrelated words, regardless of the attentional strategy in operation.

EXPERIMENT I

Method

Apparatus and materials

Twelve line-drawings were taken from a children's reading book. The drawings had similar horizontal boundaries, subtending a visual angle of approximately 2 deg when viewed through the Scientific Prototype (Model GB) tachistoscope used in the experiment. The vertical boundaries varied between 1 and 2 deg. The drawings were quite unambiguous, and represented a bird, a boat, a car, a cat, a chair, a duck, a fish, a flag, a house, a lemon, a pan and a train. Five copies of each of these drawings were taken (for the five experimental treatments of the experiment), and were pasted to white tachistoscope cards, off-centre and to the left of the card as viewed by the subjects. For each drawing a semantically related word was selected such that synonyms were excluded, but converging associates and words with high sequential dependency (in either direction) formed the set of pairs. Thus, for the picture of a chair the word TABLE was used, for the boat the word SAIL was used, and for lemon JUICE was used. The related words were between four and seven letters in length. For each related word another word of similar frequency and length was selected, but this word had no obvious semantic association with the picture. In addition to these collections of related and unrelated words, two collections of non-words were compiled. From the lists provided by Hirata & Bryden (1971) a 0-order of approximation to English and a fourth-order approximation was selected for each picture. The lengths of the approximations corresponded to the lengths of the words tied to each picture. The two words and two non-words for each picture were printed on separate tachistoscope cards and to the right of the appropriate picture as viewed by the subject. Letters were printed using 18 pt Letraset, sheet 287, in upper-case. With eye fixation in the centre of the picture the outside letter of the longest word subtended 3½ deg from the centre of fixation, and the first letter of each word subtended 1½ deg from the point of fixation. In addition to these conditions, in which a picture appeared on the same card as a word or non-word, in a fifth condition the picture was presented alone.

Onset of the display of a test card in the tachistoscope started a millisecond timer which was in turn stopped by an utterance spoken by the subject into a microphone which triggered a voice switch. The microphone was suspended from the viewing hood of the tachistoscope and was close to the subject's throat.

Subjects

Twelve students (seven females and five males) from the University of Waterloo participated in this experiment on a volunteer basis. All subjects claimed to be right-handed, but were not tested for laterality.

Procedure

Subjects were tested individually with a repeated measures design. The five experimental conditions (related word; unrelated word; 0-order approximation; fourth-order approximation; no word presented), and the 12 cards (different pictures) used in each condition were randomized in their order of presentation for each subject. The 60 trials of the experiment were presented without a break. A single trial consisted of the experimenter uttering a warning of the imminence of a stimulus, followed rapidly by the picture and the word being shown for a duration of 60 msec. The response required of the subject was the name of the picture spoken as quickly as possible after exposure.

Prior to the experimental trials the subjects were familiarized with the 12 drawings and, to remove any uncertainty, with the names which were required. Fifteen practice trials were also presented to each subject prior to testing.

The pre-test and post-test fields of the tachistoscope were both illuminated to the same intensity as the test field, and two horizontal lines in these fields indicated to the subjects the upper and lower boundaries of the set of drawings, together with their horizontal extents. Subjects

Table 1. *Focused attention. Mean response latencies (msec) in the picture-naming task of Expt. I, as a function of the type of unattended stimulus*

Unattended stimulus ...	None	0-order approximation to English	Fourth-order approximation to English	Unrelated word	Related word
Mean	581	594	591	593	617
S.D.	53	51	55	58	52

were instructed to fixate between the lines on hearing the warning before each trial, and to attend only to the picture-naming task. Subjects were informed of the presence of occasional words during the presentation of the pictures, but that they were to ignore this source of distraction. Reception of the words was never tested during the pre-experimental or experimental periods.

The latency of the naming the picture on each trial was recorded. Trials on which erroneous responses were offered were repeated later in the sequence of trials. Accuracy of response was emphasized to the subjects.

Results

Errors were recorded on less than 3 per cent of the trials, and were distributed evenly across conditions. The mean response latencies within each of the five conditions are presented in Table 1. The twelve latencies from each subject within each condition were meaned and submitted to an analysis of variance.

The effect of treatments upon the naming response was indicated to be significant ($F = 9{\cdot}1$, d.f. = 4, 44, $P < 0{\cdot}001$), and when individual words were submitted in this analysis as a non-random factor, as suggested by Clark (1973), the quasi-F ratio obtained indicated that the effect may be generalized to comparable sets of words ($F' = 3{\cdot}9$, d.f. = 4, 55, $P < 0{\cdot}01$). The five experimental treatments were compared individually by use of the Newman–Keuls test. The only difference revealed by these comparisons was that the presence of a semantically related word prolonged the latency of response to the picture over the other four treatments ($P < 0{\cdot}01$), wherein there were no differences between treatments.

Discussion

When subjects are able to focus their attention upon the location of a briefly presented stimulus before it is presented, then the printed stimulus most likely to interfere with the attended task is one which is related semantically to the picture which comes into view. This is clearly not a distraction effect because meaningless strings of letters caused no increase in the naming latency in comparison with the treatment in which no irrelevant stimulus was presented. The problem is to decide how the meaning of an unattended stimulus can influence performance towards the attended stimulus. It may be the case that all non-words and words in the visual field excite their representations in memory and that the response to the picture, or recognition of it, were disturbed by the recognition of the word. This interference may occur during picture recognition or picture name generation. Alternatively, the recognition thresholds of related words may have been lowered by the presence of contextual information provided by the recognition of the pictures. The threshold for an unrelated word would not be so affected. Once the threshold has been lowered and a semantically related word recognized the word may then interfere with the naming response. It should be noted that activation of a representation of a word is not necessarily synonymous with subjective awareness of that word. The threshold for recognition may be exceeded independently of the threshold for conscious awareness (Erdelyi, 1974; Underwood, 1976). This conclusion is necessary to account for the effects reported by Lewis (1970) and others in which subjects' performances were influenced by the meanings of words which they claimed not to have heard. Further discussion of the implications of the present results will be postponed until the following experiment has been described.

The second experiment extends the investigation of the effects of unattended words upon naming performances to a situation in which subjects employ a divided attention strategy prior to the stimulus presentation.

EXPERIMENT II

Method

Apparatus and materials

The 12 line-drawings used in the previous experiment were again used here. Eight copies of each drawing were made and four of the copies pasted to the left of a point in the centre of each tachistoscope card, and four to the right of this point. This arrangement of pictures around a fixation point in the centre of the visual field serves the presentation of the picture in the subjects' left and right visual fields. The four cards presented in these two treatments pertained to four treatments in which either (i) no word was presented with the picture; (ii) a related word was presented; (iii) an unrelated word was presented; or (iv) in which the word corresponded to the name of the picture. The unrelated words used in this experiment were words used in other trials of the experiment as words related to the picture. Thus each word appeared in two treatments, once as the related word for a picture and once as the unrelated word for another picture. This balancing procedure eliminates the effects of specific words. When subjects fixated the centre of the tachistoscopic visual field the outermost part of the widest stimulus subtended 3 deg from the point of fixation. The nearest part of either stimulus subtended 1 deg from the point of fixation.

The apparatus used was that employed in the previous experiment. The pre-test and post-test fields of the tachistoscope were illuminated and blank except for a small dot placed to coincide with the midpoint of the picture and word display.

Subjects

Twelve subjects (five females and seven males) from the University of Waterloo participated in this experiment on a volunteer basis. None had participated in the previous experiment, and all claimed to be right-handed.

Procedure

The procedure was essentially similar to that of the first experiment, with each subject being tested individually, and the 12 trials in each of the eight conditions (two levels of the visual field factor, and four levels of the unattended stimulus factor) were completely randomized in their order of presentation. The conditions were not presented as blocks of trials, but the order of the 96 cards was randomized for each subject.

The subjects were first familiarized with the set of pictures and the names which were required as the response to each. Subjects were instructed to fixate upon the black dot in the centre of the field upon hearing the warning at the start of each trial, and to report the name of the picture as soon as possible after exposure. The exposure duration was 60 msec. With the randomization procedure used subjects could never be certain as to the spatial location of the picture on any trial.

Subjects were instructed to ignore the words which sometimes accompanied the pictures, and the reception of these words was never tested. Other aspects of the procedure, including the practice and error correction procedures were similar to those described for Expt. I.

Table 2. *Divided attention. Mean response latencies (msec) in the picture-naming task of Expt. II, as a function of the type of unattended stimulus and of the field of projection of the picture*

Unattended stimulus …	None	Same word	Unrelated word	Related word
Picture in left visual field				
Mean	611	626	753	716
S.D.	101	132	127	117
Picture in right visual field				
Mean	609	629	672	657
S.D.	128	127	123	122

Results

Errors were recorded on less than 4 per cent of the trials, and were distributed evenly across conditions. The mean naming latencies within each of the four conditions for presentation in each visual field are given in Table 2. The means of the 12 latencies obtained from each subject in each of the eight conditions were submitted to an analysis of variance, which indicated that both of the main treatments were effective.

Responses to pictures in the right visual field were faster than to those in the left visual field ($F = 51{\cdot}5$, d.f. = 1, 11, $P < 0{\cdot}001$), and the irrelevant words also affected latency ($F = 15{\cdot}2$, d.f. = 3, 33, $P < 0{\cdot}001$). Clark's (1973) quasi-F indicated the generality of the effect ($F' = 5{\cdot}1$, d.f. = 4, 44, $P < 0{\cdot}01$). The interaction between the two main factors ($F = 10{\cdot}3$, d.f. = 3, 33, $P < 0{\cdot}01$) indicates the greater effect of related and unrelated words when these words are projected to subjects' left cerebral hemispheres. This generalized supposition was tested with two Newman–Keuls tests, one between the four conditions in which the picture was in the left visual field and one between the conditions of right visual field presentation.

When the picture was projected to subjects' right hemispheres the Newman–Keuls analysis indicated no difference between the treatments where no word was presented, and where the word was the same as the response offered on that trial. Presentation of a related word slowed the naming response in comparison with these two treatments (significant at the $P < 0{\cdot}01$ level), and presentation of an unrelated word increased the naming latency even further ($P < 0{\cdot}01$). The pattern changes slightly when considering the test applied to the four treatments when the picture was projected to subjects' left hemispheres. Presenting the same word had a similar effect as presenting no word at all, and both the related-word ($P < 0{\cdot}01$) and the unrelated-word ($P < 0{\cdot}01$) treatments differed from the same-word treatment, but there was no difference between the effects of the related and unrelated words by this test.

Discussion

A potent factor in this experiment was the visual field of presentation of the task-irrelevant word; when the word was projected to the left hemisphere it caused a greater interference effect than when projected to the right hemisphere. Words and single letters are known to be identified more accurately when exposed in the right visual field (Heron, 1957; Bryden, 1965). It is not clear from the present data whether the increased interference effect of words in the right visual field over words in the left visual field is a function of superior contact with the lexicon, or of increased output interference following superior contact with the speech production centres located in the left cerebral hemisphere. Indeed, it is not even clear that this distinction can be maintained. If the lexical store of verbal representations is an integral part of the speech production system then contacting the lexicon may necessarily lead to the production of a phonemic representation, and response competition. The relatively low error rate in this experiment works against this hypothesis, in that observers might have been expected to utter more of the printed words in error had their phonemic representations been available whenever their meanings were available. Reading is notoriously easier than naming (Fraisse, 1969), and yet the presentation of the naming response in the form of a word (i.e. in the same-word condition) did not facilitate the naming response whether presented to the left or the right hemisphere. This would suggest that the phonemic representation was not available when attention was directed towards the picture, or that the response was not offered until the name had been generated and correlated with the name available from processing of the word. Presentation and rapid processing of the word in the same-word condition did not facilitate processing of the picture and generation of the picture name, and this is where the present effect differs from the interference-facilitation effect of the Stroop test. Presentation of a colour name in that same colour of ink facilitated naming of the colour of ink in comparison with trials where the colour name and colour of ink did not match (Hintzman *et al.* 1972), and where the name was a non-word or an unrelated word. Whereas attention may prevent the generation of the phonemic representation of a task-irrelevant word in the present experiment it does not prevent facilitation or interference with the Stroop words.

The related and unrelated words had differential effects according to the visual field in which they were presented, and so the task-irrelevant words were not processed independently of the pictures. Both classes of words caused more naming interference than did the same word as the picture, indicating that the effect is not a function of a distracting visual stimulus whose meaning is irrelevant. Any distracting effect of the same-word stimulus was balanced by the facilitatory effect of the word coinciding with the name being sought for the picture, but this facilitation was not available to such an extent for related and unrelated words. Relative facilitation may also account for the difference between the effects of related and unrelated words (for right visual field presentations). The interference caused by the processing of these words may have been counteracted slightly when the word was part of the semantic network associated with the picture name being sought. If the printed word excited a branch of the association network being used for the primary task, then the interference might be expected to be less than when a different branch is excited. Analysing two related items would then be held to be a more efficient process than analysing two unrelated items given that both items must be analysed each time. Meyer & Schvaneveldt (1971) also provide data relevant to this explanation. When responding to two letter-strings as both being words or otherwise, subjects found that task easier (for the case of two words) if the words were semantically related. By the spreading-excitation model considered by Meyer & Schvaneveldt the retrieval of our picture name would facilitate the retrieval of words stored in nearby locations.

Words in 'nearby locations' would be expected to be those which are semantically associated (Collins & Quillian, 1969).

One problem with this theory of facilitation-interference due to excitation of lexical association networks is raised by the data from Expt. I, and this will be explored in the following section.

GENERAL DISCUSSION

The interference effect in the second experiment was greater for unrelated words presented to the left cerebral hemisphere than it was for related words, but in the first experiment (with left hemisphere projection) a word unrelated to the picture caused no interference with the latency of the naming response whilst a related word did cause a slight retardation of that latency. This comparison demands a revision of the facilitation-interference hypothesis of simultaneous processing, in that related words should cause some interference with the naming task in both situations, but not as much as unrelated words. At this point it is necessary to take a more general view of the processing strategies operating in the two situations. In the first experiment the subjects might be expected to attend closely to the spatial location of the impending stimulus, and have no necessity to attend to any aspect of the irrelevant visual field. In the second experiment attention might be expected to be divided between the two spatial locations, and processing of the irrelevant stimulus terminated volitionally only when the observer is able to decide which of the two stimuli is the picture. If that decision can be made prior to the extraction of meaning from the printed symbols then no semantic interference will be recorded, but the present data indicate that attention cannot be switched away from the word so quickly, or that the word is processed semantically regardless of the direction of attention. However, the two experiments differ fundamentally in the attentional strategies which are appropriate for performance. When attention is focused then only slight semantic interference is apparent, and this is restricted to an effect of a related stimulus. When attention is divided a different pattern is apparent, and the interference effect is more general. These data are consistent with a conclusion taken from an experiment reported by Treisman, Squire & Green (1974), which was a replication of the Lewis (1970) dichotic listening experiment. Whereas Lewis found that unattended synonyms retarded the latency of shadowing simultaneously presented words, Treisman *et al.* found that this interference effect was restricted to the early pairs of items in each dichotic list, when attentional selectivity is arguably poor. Focused attention may be a state which takes a few seconds to achieve when listening to dichotic stimuli, and if a potentially interfering stimulus is presented during this interval then it would be more effective than at other times. With divided attention we accept more information from the environment, however transient this state may be, and however aware of this information we may be. By allowing more information to be processed in this way the irrelevant words affect the ongoing activities with an interference-facilitation interaction. Posner & Snyder's (1975) views of the function of attention appear to be inadequate, for they do not take into account the different affects of attentional strategies. When attention was divided and relatively more stimulus information accepted, the thresholds of recognition of irrelevant words were exceeded, and these recognition responses consequently affected the processing of the picture. When the recognized word was a part of the association network then relatively little interference was experienced than when the word was semantically unrelated and could offer no facilitation. In the focused attention situation relatively less stimulus information was accepted and thresholds not exceeded; the unattended stimulus might then be described as being attenuated, following Treisman (1960). With a semantically related word the recognition of the picture acted to lower the recognition threshold of the word, and only then did activation of the logogen of the word interfere with the processing of the picture name.

The foregoing speculations imply that two semantically related stimuli will be recognized regardless of the direction of attention and provided that one of them is recognized. This is analogous to recognition thresholds of visually presented words being lowered by prior contextual information, as described by Tulving & Gold (1963). When attention is less focused then semantically unrelated words may also be recognized, but via a different route: their thresholds are affected only by stimulus information. The recognition of such words which cannot facilitate the performance of the ongoing activity will necessarily lead to greater interference than would the recognition of semantically related words. Activation of a lexical representation does not mean that the subject would necessarily be aware of that recognition or the consequent effect. This model finds support from studies of the reading process, in which unattended words have been found to influence the comprehension of an attended passage (Willows, 1974), and Kolers (1970) has argued that skilled readers are able to extract the meaning of text without directed reference to the appearances or sounds of individual words. Words can activate their logogens without phonological mediation, and LaBerge & Samuels (1974) also propose a model which dispenses with the necessity for attention to words prior to their recognition.

Attention may not be required for word recognition, provided that the thresholds of words are lowered by current contextual information. When operating a strategy of focused attention only contextually related words may be recognized, and this will have a general effect, when reading prose, of providing more information about the idea being comprehended. An alternative mode of operation is to employ a divided attention strategy, in which contextually improbable words may affect comprehension and generally bias the interpretation of perceptual ambiguities. The function of attention might then be found in the resolution of ambiguity and in the elaboration of our schemata, as Wundt suggested.

These experiments were conducted while the author was on leave at the University of Waterloo, and were aided by support from a National Research Council of Canada grant (no. A-95) to M. P. Bryden, which is gratefully acknowledged.

Current address: Department of Psychology, University of Nottingham, Nottingham, NG7 2RD.

REFERENCES

ANDERSON, B. F. (1975). *Cognitive Psychology*. New York: Academic Press.

BRADSHAW, J. L. (1974). Peripherally presented and unreported words may bias the perceived meaning of a centrally fixated homograph. *J. exp. Psychol.* **103**, 1200–1202.

BROADBENT, D. E. (1958). *Perception and Communication*. London: Pergamon Press.

BRYDEN, M. P. (1965). Tachistoscopic recognition, handedness, and cerebral dominance. *Neuropsychologia* **3**, 1–8.

CLARK, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *J. verb. Learn. verb. Behav.* **12**, 335–359.

COLLINS, A. M. & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *J. verb. Learn. verb. Behav.* **8**, 240–247.

CONRAD, C. (1974). Content effects in sentence comprehension: A study of the subjective lexicon. *Mem. Cog.* **2**, 130–138.

CORTEEN, R. S. & WOOD, B. (1972). Automatic reponses to shock associated words in an unattended channel. *J. exp. Psychol.* **94**, 308–313.

DEUTSCH, J. A. & DEUTSCH, D. (1963). Attention: Some theoretical considerations. *Psychol. Rev.* **70**, 80–90.

ERDELYI, M. H. (1974). A new look at the new look: Perceptual defence and vigilance. *Psychol. Rev.* **81**, 1–25.

FRAISSE, P. (1969). Why is naming longer than reading? *Acta Psychol.* **30**, 96–103.

HERON, W. (1957). Perception as a function of retinal locus and attention. *Am. J. Psychol.* **70**, 38–48.

HINTZMAN, D. L., CARRE, F. A., ESKRIDGE, V. L., OWENS, A. M., SHAFF, S. S. & SPARKS, M. E. (1972). 'Stroop' effect: Input or output phenomenon? *J. exp. Psychol.* **95**, 458–459.

HIRATA, K. & BRYDEN, M. P. (1971). Tables of letter sequences varying in order of approximation to English. *Psychon. Sci.* **25**, 322–324.

JACOBSON, J. Z. (1973). Effects of association upon masking and reading latency. *Can. J. Psychol.* **27**, 58–69.

KOLERS, P. A. (1970). Three stages in reading. In H. Levin & J. Williams (eds), *Basic Studies in Reading*, pp. 90–118. New York: Basic Books.

LABERGE, D. & SAMUELS, S. J. (1974). Toward a theory of automatic information processing in reading. *Cog. Psychol.* **6**, 293–323.

LEWIS, J. L. (1970). Semantic processing of unattended messages using dichotic listening. *J. exp. Psychol.* **85**, 225–228.

MACKAY, D. G. (1973). Aspects of the theory of comprehension, memory and attention. *Q. Jl exp. Psychol.* **25**, 22–40.

MEYER, D. E. & SCHVANEVELDT, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *J. exp. Psychol.* **90**, 227–234.

MORTON, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* **76**, 165–178.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

NORMAN, D. A. (1968). Toward a theory of memory and attention. *Psychol. Rev.* **75**, 522–536.

POSNER, M. I. & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. Solso (ed.), *Information Processing and Cognition: The Loyola Symposium*. Potomac, Maryland: L. Erlbaum Associates.

SMITH, M. C. & GROEN, M. (1974). Evidence for semantic analysis of unattended verbal items. *J. exp. Psychol.* **102**, 595–603.

TREISMAN, A. M. (1960). Contextual cues in selective listening. *Q. Jl exp. Psychol.* **12**, 242–248.

TREISMAN, A. M. (1969). Strategies and models of selective attention. *Psychol. Rev.* **76**, 282–299.

TREISMAN, A. M. & GEFFEN, G. (1967). Selective attention: perception or response? *Q. Jl exp. Psychol.* **19**, 1–17.

TREISMAN, A. M., SQUIRE, R. & GREEN, J. (1974). Semantic processing in dichotic listening: A replication. *Mem. Cog.* **2**, 641–646.

TULVING, E. & GOLD, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *J. exp. Psychol.* **66**, 319–327.

UNDERWOOD, G. (1972). Response organization in attention control and a perceptual serial position effect. *Q. Jl exp. Psychol.* **24**, 340–351.

UNDERWOOD, G. (1973). Control of selective attention and interference of processing in memory. *J. exp. Psychol.* **99**, 28–34.

UNDERWOOD, G. (1976). *Attention and Memory*. Oxford: Pergamon Press.

UNDERWOOD, G. & MORAY, N. (1971). Shadowing and monitoring for selective attention. *Q. Jl exp. Psychol.* **23**, 284–295.

WILLOWS, D. M. (1974). Reading between the lines: Selective attention in good and poor readers. *Child Dev.* **45**, 408–415.

WUNDT, W. (1897). *Grundriss der Psychologie.* Leipzig: Engelmann.

(*Manuscript received* 1 *April* 1975; *revised manuscript received* 27 *May* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 339–351
Printed in Great Britain

SEMANTIC CODING VERSUS THE STIMULUS SUFFIX

BY D. SALTER, G. SPRINGER AND L. BOLTON
Department of Psychology, University of Newcastle-upon-Tyne

The stimulus suffix effect is investigated when it follows an auditory memory list in which rated meaningfulness was manipulated at the final serial position. A verbal suffix disrupts the terminal list item compared with a noise suffix, but rated meaningfulness affects recall performance significantly despite the presence of a verbal suffix. With ordered recall, there is some evidence that a verbal suffix disrupts items rated low in meaningfulness to a greater extent (Expt. I); this interaction does not show when serial order recall is stipulated (Expt. II). The effect of rated meaningfulness with a verbal stimulus suffix eliminates a model in which information about the final item is retrieved exclusively from precategorical acoustic storage. The paper discusses two propositions about the preliminary stages of acoustic analysis and encoding in the absence of focal attention.

This paper examines the stimulus suffix effect in the context of a recent model formulated by Morton (1968, 1970), which describes explicitly a number of storage systems for different types of coding.

The model, which is schematically represented in Fig. 1, identifies an input buffer process, that operates through modality specific analysers to extract prelinguistic features of language stimuli. Our concern in this paper is with the auditory analysers of *precategorical acoustic storage* (PAS) whereby features of the verbal input are held initially in a 'veridical' form.

The input buffer process is distinguished from an output buffer process, labelled the *response buffer*, in which articulatory programs are serially initiated. Linking these two buffer processes is a lexically organized interface of *logogen* units, which are connected to a longer-term storage system: the *cognitive system*. As partial information about the stimulus properties of a word are transferred from the input analyser to the *logogen system*, incremental activity for a particular unit mounts toward some critical value, at which an articulatory code becomes available at the output buffer. The storage characteristics of the response buffer may be likened to the notion of primary memory (see Craik, 1971, for review). At this stage the coded representation may be converted either to an overt vocal response, or to a process of covert rehearsal by recycling named features back to the logogen system. The logogen units themselves are not regarded as a mode of stimulus representation, but are conceived as counting indices for a sufficiency of evidence required to convert one kind of representation to another. When sufficient stimulus information is registered at any given logogen, conversion to semantic coding automatically takes place; Morton (1970) suggests that a lower threshold of logogen activity is required for registration in the cognitive system than that threshold required for conversion to an articulatory code at the response buffer. Evidence for registration without awareness is discussed by Coltheart (1972, pp. 62–85) and might be cited as illustrating the autonomous conversion of information into a semantic code.

The recency effect, which is observed in the immediate serial recall of a spoken memory string, has been attributed by Crowder & Morton (1969) to the relatively short-lived acoustic representation retained in PAS. Information in PAS by definition

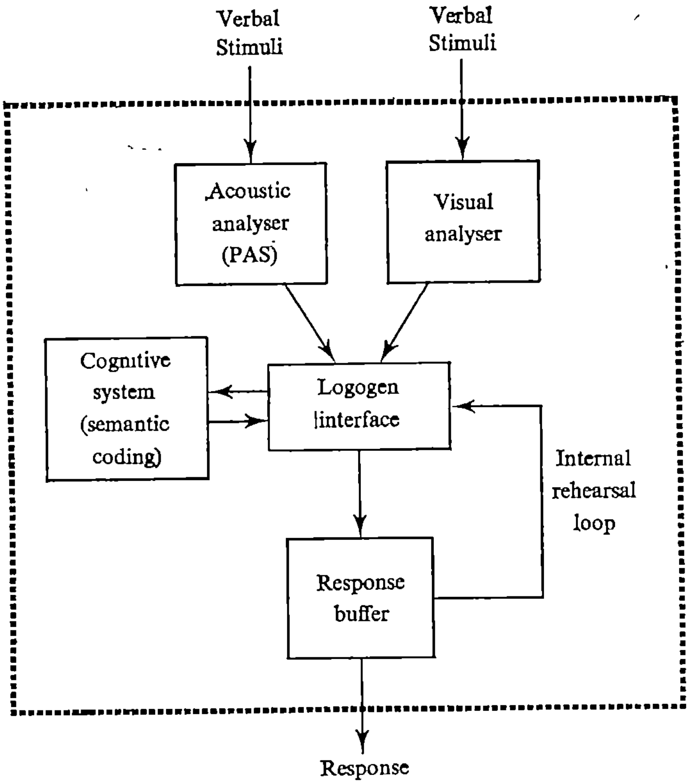


Fig. 1. Schematic representation of the logogen model (after Morton, 1970).

is coded in terms of physical features and is context independent. Preliminary estimates of the duration for this information suggested a period of about 2 sec (Crowder, 1969; Darwin, Turvey & Crowder, 1972), though more recent evidence has extended this time (Crowder, 1971; Routh & Mayes, 1974; Cole, Coltheart & Allard, 1974). This temporary source of acoustic information benefits recall for the last one or two items of the memory list, but this advantage is nullified by the addition to the memory string of one extra item, even though the subject is instructed to ignore it (Crowder & Morton, 1969; Crowder & Raeburn, 1970). The effect of this *stimulus suffix* is to produce marked but selective impairment in the recall for the terminal items of the memory string. This effect can be attenuated progressively by changing the physical characteristics, which identify the suffix, from those which identify the memory list. (For a recent experimental review of the phenomenon, see Morton, Crowder & Prussin, 1971.) Crowder & Morton (1969) have suggested that the effect to the suffix is to overwrite or to displace existing PAS information which is crucial for the recall of terminal items.

Morton *et al.* (1971, Expt. V) report that the size of the suffix effect is not influenced by varying the semantic similarity of the suffix and the list items. Given that the disruptive effect of the suffix is on the final item predominantly, one interpretation of this evidence might reasonably be that final item information is displaced by a verbal suffix from PAS before its meaning is analysed by the cognitive system. Yet elsewhere Morton (1970, p. 214) specifies that input is analysed and resynthesized at presentation via the cognitive system, before the response is made available in the response buffer. In this case a verbal suffix would not only overwrite the existing representation in PAS, but it would also interfere to some extent with a second representation, held simultaneously, in the response buffer. Nevertheless the cognitive system would retain relevant information, which could be used for recall subsequently.

In either case a verbal suffix disrupts performance at the final serial position compared with the corresponding noise condition, but its effect could be attenuated in the second case if the final item was highly meaningful. The novel aspect of our experiments is that the level of meaningfulness of list items has been systematically controlled. In particular the rated meaningfulness of the final list item is varied as a method of distinguishing between the two alternatives outlined above. In the first case, the difference in meaning for the final item should make no difference to recall performance since the information is displaced from PAS before the semantic attributes of its meaning content become effective. In the second case, the final item is extensively coded at presentation so that its rated meaningfulness should affect recall performance even though the list is followed by a verbal suffix, which overwrites existing information in PAS.

In the first experiment we wished to see whether a systematic effect from the last item obtained. Words and paralogs, employed as list items, were recalled in the correct serial position but in any order of recall. The usual strategy with ordered recall is for later items to be output first. This recall method does not therefore disadvantage paralogs, with low M rating, by lack of response availability to the same extent as does serial ordered recall. Testing with both noise and verbal suffix conditions enabled us to see whether a systematic effect for the final list item would show both across and within each type of suffix condition.

EXPERIMENT I

Method

Materials

Stimulus lists were selected from CVCVC items, scale-rated for meaningfulness (M) by Locascio & Ley (1972). These items were assigned to two blocks of eight lists, seven items per list, with the following restrictions: (*a*) each item had a scale-rated meaningfulness greater than 1·50 but less than 3·50; and (*b*) each list contained one item from each of the following categories of meaningfulness:

$$1{\cdot}50 < M_1 < 2{\cdot}00,\ 2{\cdot}01 < M_2 < 2{\cdot}25,\ 2{\cdot}26 < M_3 < 2{\cdot}50,\ 2{\cdot}51 < M_4 < 2{\cdot}75,$$
$$2{\cdot}76 < M_5 < 3{\cdot}00,\ 3{\cdot}01 < M_6 < 3{\cdot}25,\ 3{\cdot}26 < M_7 < 3{\cdot}50.$$

These seven levels of rated meaningfulness were distributed as equally as it was possible, given the stimulus constraints, across the serial positions of the recall lists.

These are termed the 'base' lists. Onto each base list an eighth item was mapped, either of very high meaning ($M > 4{\cdot}50$) or of a very low scale-rated meaning ($M < 0{\cdot}60$) such that the final items on alternate pairs of lists were of different meaning values. (These final items are given in an Appendix.)

A CVCVC suffix (the word *begin*) or a non-suffix control (a burst of 'white noise') were also mapped onto each list (Crowder & Raeburn, 1970) such that the suffix and the non-suffix control occurred alternately on lists with either high or low M-rated eighth items.

Thus four conditions varying the meaningfulness of the eighth item and the type of suffix were mapped onto the base lists. Four sets of stimulus lists were generated; each set consisted of a different mapping of the four condition treatments onto the base lists.

The stimulus lists were recorded in an English male voice at the rate of one item per 0·75 sec using a twin-channel Revox recorder. The suffix or white noise was recorded as if it were a ninth list item with the same inter-stimulus interval. Each list was preceded by the word *Ready*. The interlist interval was 20 sec. Each set of eight lists was recorded separately.

Subjects

Twenty subjects were tested; 13 were men. They were all undergraduates of the University of Newcastle. The mean age was 21 years with an age range of 18–33 years. All the subjects were unpaid volunteers and were naive with respect to this particular experimental procedure. Subjects were tested in groups of five.

Procedure

Four groups were tested in the same room and at the same time on different days. Each group received a different set of stimulus lists. Four practice lists were given prior to the experimental trials, followed by the two blocks of eight lists. Testing lasted approximately 30 min with a brief rest pause at the half-way point.

The instructions were that each list consisted of eight words, some genuine and some nonsense words, and that lists would end with either the instruction *Begin* or a burst of white noise. The occurrence of either of these signals was a cue to write down as many words as could be remembered, but excluding the word *Begin*. Booklets were prepared with one row of box spaces per stimulus list per page. Responses were to be written in the appropriate box space. Subjects were told to approximate their spelling of the words to the sound as closely as possible. At the bottom of every instruction sheet, an example of a full response list was given. This also gave some indication of the nature of the stimulus items.

Results

The criterion for scoring specified that if the item was correctly spelt it was scored 2, if almost correct it was scored 1 given that the response was in the appropriate serial position on the recall sheet. Incorrect responses scored zero. Some of the less meaningful items were misheard or transformed into meaningful words. Thus NEROL ($M = 0{\cdot}52$) became NEURAL or NEURON; KETEL ($M = 1{\cdot}56$) became KETTLE; TORUS ($M = 2{\cdot}58$) became TAURUS. To maintain a constant scoring criterion, the data were rescored by a second, neutral observer. The few discrepancies which arose from this procedure were fairly easily resolved. Serial position curves for these data are given in Fig. 2. Each point averaged over subjects is based on 80 observations.

Analysis of variance was made over the last serial position. This tested the within-subject variables: two levels of suffix, and two levels of meaning. The analysis assumed fixed effects. The raw data were transformed by the statistic $\sqrt{(X+0{\cdot}5)}$. The two main effects only were significant: verbal suffix/noise suffix $F = 19{\cdot}33$, d.f. = 1, 16, $P < 0{\cdot}001$; high/low rated meaning $F = 22{\cdot}0$, d.f. = 1, 16, $P < 0{\cdot}001$. The interaction between suffix and meaning on the transformed data was not significant $F = 0{\cdot}80$.

Wilcoxon tests were also made on the untransformed data for the last serial position. These tests confirmed that recall was superior with a noise suffix compared with a verbal suffix for high M-rated items at the $P < 0{\cdot}05$ level, and for the low M-rated items at the $P < 0{\cdot}01$ level (two-tailed). The untransformed data also show some evidence of interaction with a larger difference between the two low M conditions

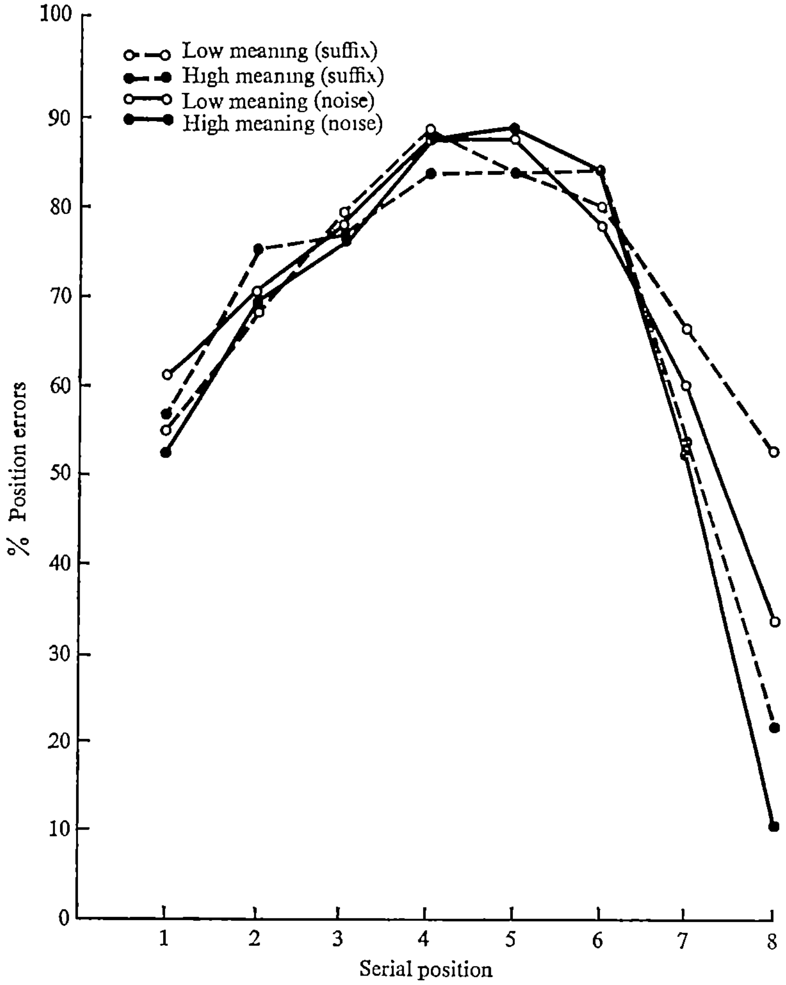


Fig. 2. Percentage errors for memory strings with a final item of high or low rated meaningfulness, followed by a suffix item or its control condition. Order of recall not stipulated.

compared with their respective high M conditions (see Fig. 2). This effect is reduced by a square root transform in the analysis of variance but shows with untransformed data. Therefore a Wilcoxon test was made between the high M versus the low M condition using a difference score (verbal suffix − noise suffix). This was significant $T = 49$, $N = 20$, $P < 0{\cdot}05$ (two-tailed).

Discussion

Recall performance for the final list item is markedly different for the two levels of rated meaning, and this is also true for conditions in which a verbal suffix is present. This latter observation thus provides some evidence that the auditory input is coded at a stage subsequent to PAS.

The question of where this differential effect for rated meaningfulness occurs is, we think, more difficult to isolate. Various explanations suggest themselves. The effect could be due to processing in the cognitive system as Morton (1970) proposes, or more simply the effect can be attributed to the lack of an appropriate response becoming available in the response buffer, or an interaction of both these causes. With ordered free recall, it is likely that the final list items are recalled first, so there is also the possibility that these items are selectively rehearsed before recall and transferred from the response buffer to the cognitive system via the rehearsal loop (see Fig. 1). The relatively poor performance over the early serial positions (Fig. 2) lends support to this interpretation, but could also be attributed to response interference. In either case, the data suggest that the later serial positions are output first.

Having established an effect for rated meaning with a verbal suffix, we then attempted to reduce these alternatives by requiring serial ordered recall. This stricter recall procedure ensures that the short-term capacity of the response buffer is fully occupied with the first few items of the memory string. (The capacity of the response buffer is estimated by Morton (1970) at about four items, which are processed serially.) We also emphasized in the instructions that subjects were to concentrate on remembering the first few items correctly, and then any others they could remember. This enabled us to see whether the rated meaning effect still occurred.

The use of CVCVC paralogs also posed difficulties for scoring. Because the spelling of responses approximated to the sound of each item, the way was opened for item replacement with words of higher meaning ratings. These items scored a zero rating, but their effect on other items in the list remains unspecified and is best avoided altogether. Untoward effects of this kind were balanced out to the extent that identical base lists were combined with different treatment conditions across subject groups, and theoretical interest in the experimental effect was also restricted to recall for the final serial position only. Nevertheless, we considered it necessary to attempt a replication of the effect in a second experiment which imposed more stringent controls on the selection of base list items, and on recall procedure.

EXPERIMENT II

Method

Materials

Two sets of eight lists were generated using trigrams in the form CVC, selected from those listed by Noble, Sutker & Jones (1968). List items were chosen from one of the following levels of scaled meaningfulness (M):

$$(1)\ 2{\cdot}50 < M < 2{\cdot}99;\ (2)\ 3{\cdot}00 < M < 3{\cdot}24;\ (3)\ 3{\cdot}25 < M < 3{\cdot}49;\ (4)\ 3{\cdot}50 < M < 3{\cdot}74;$$
$$(5)\ 3{\cdot}75 < M < 3{\cdot}99;\ (6)\ 4{\cdot}00 < M < 4{\cdot}24;\ (7)\ 4{\cdot}25 < M < 4{\cdot}49.$$

The ordering of items was randomized within the memory string with the constraint that M values were distributed as evenly as possible across all lists.

An eighth item, chosen from the M scale of meaning at high ($M > 5{\cdot}25$) or at low ($M < 1{\cdot}75$), was then randomly assigned to each list at the last serial position so that there was an equal number of lists with high and with low meaningful eighth items (see Appendix). The word *Begin* to function as a suffix, or a burst of white noise to act as a suffix control, was then mapped onto each list. This mapping procedure produced four combinations for eighth item plus suffix as the experimental conditions: high meaning/suffix, high meaning/noise, low meaning/suffix, and low meaning/noise.

The four versions of the stimulus lists thus produced were recorded in a male voice using a twin-channel Revox recorder. Items were recorded at a faster rate than previously of two items per second with the suffix or its control (a noise burst of the same duration) being treated as a ninth item. Lists were presented at the rate of three per minute so that the interlist recall interval was approximately 14 sec. There were four different orders of presentation.

Subjects

Eleven women and 13 men from the University of Newcastle participated. The age range was 19–25 years with a mean of 21 years. All subjects were unpaid volunteers and new to the experimental procedure. These subjects were randomly assigned to the four presentation orders, and tested in groups of three in an experimental session which lasted about 30 min.

Procedure

The gist of the subjects' instructions was as follows. Lists of eight items were to be presented with each item in the form of a consonant–vowel–consonant. These items were a mixture of words and nonsense syllables. Immediately following each list, there was a burst of static noise or the word *Begin*. These were signals to begin writing down the memory list. Subjects were reminded that their responses were not to include either the instruction *Begin*, or the word *Ready* which preceded all the lists.

It was emphasized that list items were *to be recalled in the order in which they had been heard.* Subjects were to write down the items from left-to-right in the box spaces provided without backtracking. They were asked to concentrate on remembering the first few words correctly and then any others they could remember, guessing if necessary. Responses were written down in a row of eight boxes on a separate page of a prepared response booklet. The instruction sheet also listed 48 items as examples, which were similar but not identical to the experimental material. This fact was pointed out to the subjects.

All groups received eight practice trials, then two sets of stimulus lists. There was a short rest pause at the half-way point. Testing was carried out at the same time of day in the same room. Stimulus lists were relayed via headphones to each subject.

Results

Items had been selected with some care so that items which were ambiguous with respect to spelling were avoided. This provision allowed the stringent scoring criterion to be used that responses to be correct were spelled exactly and were located in the appropriate serial position. Errors were largely intralist confusions and misperceptions, though some confabulations also occurred. Serial position curves for the error data are presented in Fig. 3. Each data point is based on 192 observations averaged over subjects.

Analysis of variance for repeated measures was carried out on the error scores for the last serial position [transformed $\sqrt{(x+0{\cdot}5)}$]. Reliable differences were found between the verbal suffix/noise suffix conditions $F = 77{\cdot}23$, d.f. $= 1, 20$, $P < 0{\cdot}001$, and between the high/low meaning value for the eight item $F = 50{\cdot}27$, d.f. $= 1, 20$, $P < 0{\cdot}001$. The interaction between suffix type and meaning value was not significant $F = 0{\cdot}02$.

A Page's L test was made on the final serial position scores across condition treatments with untransformed data. This analysis tested the original proposal that high M-rated items would be more favoured in recall and, subsidiary to this effect, that terminal items would be less disrupted by a noise burst than by a verbal suffix. There would be less justification for this hypothesis in this order of priority had it not been shown that changing the alphanumeric category of the final item of a list overrides the suffix effect (Salter, 1975), together with the results from the preceding experiment. The hypothesis was strongly confirmed $L = 644$, $M = 24$, $n = 4$, $P < 0{\cdot}001$. The relative advantage of high versus low M-rated items for the two intermediate treatments (Fig. 3) is less well defined with shorter CVC stimulus items compared with those used in the previous experiment (see Appendix).

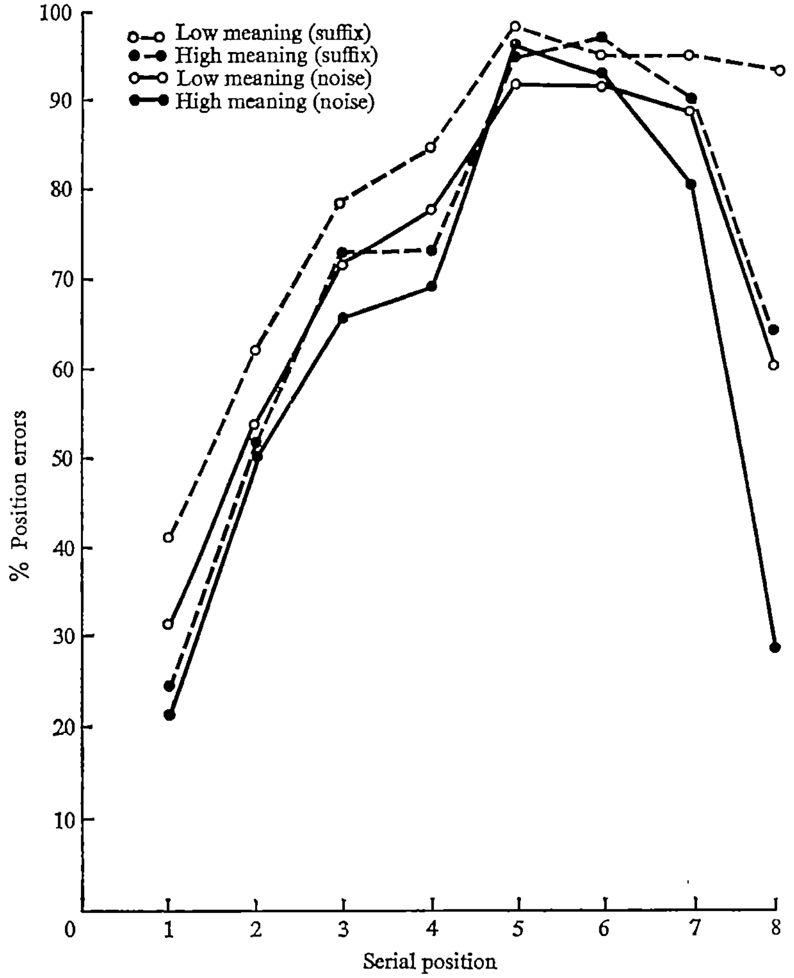


Fig. 3. Percentage errors for memory strings with a final item of high or low rated meaningfulness, followed by a suffix item or its control condition. Left-to-right order of recall required.

Final serial position scores for the verbal suffix condition were reliably different from the noise suffix condition (Wilcoxon; $P < 0{\cdot}01$, two-tailed) for the high M-rated items, and for the low M-rated items.

Discussion

Even with the more stringent procedural requirements these results confirm the difference between the two levels of rated meaning obtained in the first experiment. A change in rated meaning at the last serial position counteracts the deleterious effect on recall of the stimulus suffix. We can infer with greater confidence that the final item is coded in the cognitive system, albeit with a cautionary proviso that rated meaningfulness has also been found to relate to other variables as well (see Underwood & Schulz, 1960). Given that semantic coding is operating here, these data support proposals that information may exist simultaneously in more than one form of coding (Peterson, 1969; Morton, 1970; Craik & Lockhart, 1972), and that information

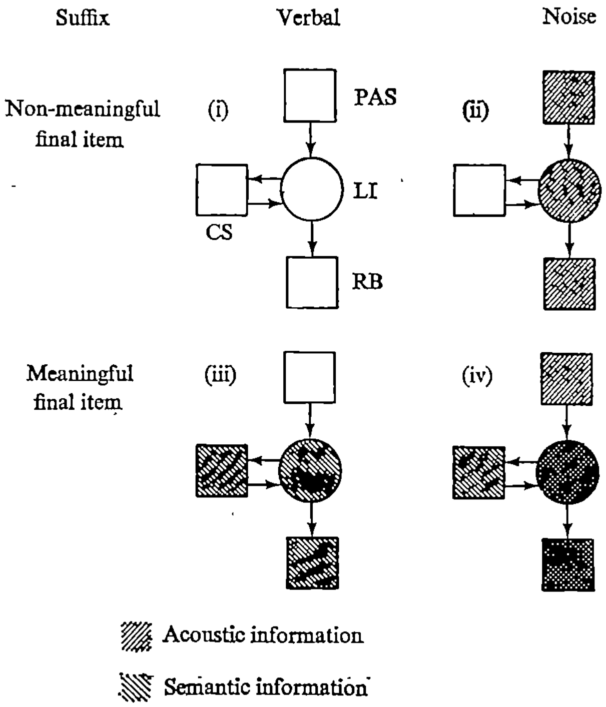


Fig. 4. A model of information processing in precategorical acoustic storage (PAS), logogen interface (LI), cognitive system (CS), and response buffer (RB) for conditions with a final recall item of high or low rated meaningfulness, and a suffix or its control.

is not 'stripped away' when it is recoded (Liberman, Mattingly & Turvey, 1972; Darwin & Baddeley, 1974). These results may also be compared with those obtained when the alphanumeric category of the final item was manipulated (Salter, 1975). In that study a change in alphanumeric category eliminated the suffix effect altogether, whereas in this second experiment the size of the suffix effect remains approximately the same at both levels of rated meaningfulness.

In the model outlined in Fig. 1 (Morton, 1970), information is transferred to the cognitive system via the logogen interface. Morton has proposed (1968, 1970) that the logogen has two thresholds. The lower threshold defines a criterion level for access to the cognitive semantic system, and the higher threshold is critical for setting up an articulatory program at the output response buffer. Data reported by Routh (1971) and Routh & Mayes (1974) suggest that this information transfer from PAS to the cognitive system is autonomous. A similar proposal of automatic 'registration' of incoming information against a trace register in permanent memory is also discussed by Salter (1973) in the context of selective attention.

Information processing in the stimulus suffix paradigm might be formulated in the following way. With no suffix, information about the last item is available for recall jointly from PAS and also from the cognitive system if the item is rated high in meaning. This is illustrated in Fig. 4 (ii) and (iv) respectively. With a suffix, the traces of acoustic information are displaced but encoded semantic information remains for correct recall if the item is rated high in meaning, 4 (iii). According to the model, little or no information is available from the cognitive system when the final item of the memory string is rated low in meaningfulness, even given that an appropriate logogen interface exists for the initial transfer to the cognitive system. Recall information about a final item rated low in meaningfulness must therefore be derived mainly from 'veridical' sources which are associated with PAS. In this latter condition the suffix is predicted to affect recall drastically, 4(i), and this is observed (Fig. 3).

An alternative model to the foregoing is one in which the suffix effect is proposed to be more central in origin. This interpretation is supported by Massaro, who states that the suffix effect is due to the interference of the subject's short-term memory for the final items of the lists (Massaro, 1972, pp. 140–141). However, if by *short-term memory*, we read *response buffer* (see Morton, 1970), then a number of objections to interference at this stage in the system can be made; specific points are mentioned in Darwin & Baddeley (1974, p. 55) and Salter (1975). Quite apart from that, the effect of rated meaningfulness in these two latest experiments also raises some difficulty for an interpretation which isolates interference to within a short-term memory store.

Our interpretation of the data wishes to emphasize the processes of perceptual encoding rather than to distinguish separate storage *components*, regarding storage as a concomitant of the coding process. Two proposals are made.

The first proposes that the non-interference from a 'noise' suffix, or some similar control, is not because this stimulus information is excluded from a specifically linguistic device, such as PAS, but because white noise is separated from the verbal items of the memory string along the encoding dimensions within the acoustic analyser. The suggestion is made on the following grounds. The attenuation of the suffix effect with physical differences between suffix and string is now well documented (Morton, 1970; Crowder, 1972). Darwin & Baddeley (1974) have shown furthermore that what was previously interpreted in terms of linguistic categories between speech sounds (Smallwood & Tromater, 1971; Crowder, 1971, 1972; Cole, 1973) can be explained parsimoniously in terms of acoustic discriminability alone. If the distinction between speech and non-speech sounds within the acoustic analyser is allowed to lapse, then the curious asymmetrical suffix effect obtained by Routh & Lifschutz (1975) can be interpreted more readily. In their study, eight memory items paired with a burst of noise and followed by a verbal suffix without noise did not show a suffix effect, but the converse condition did. We suggest that an acoustic analyser which is accessible to a wider spectrum of input than the existing logogen model allows seems necessary to explain this evidence.

Our second proposal is that incoming information automatically activates representations of similar inputs in permanent memory without this *necessarily* entailing access via a limited capacity processor, which is here equated with focal attention. When a representation is activated in this way, the probability of its accurate location and the retrieval of the *registration* to focal attention is severely reduced if another item rapidly follows. Although the rate of stimulus presentation before the suffix allows adequate time for perceptual registration and identification, the suffix seems to prevent further processing of the item immediately before it. Routh & Lifschutz (1975) have compared the function of the suffix in this respect with Turvey's (1973) analysis of central visual backward masking and refer to Kolers'

(1968) clerk–customer analogy to describe the process. The excess time devoted to each customer is reduced relative to the number waiting to be served. We might add, however, that if the clerk remains almost entirely preoccupied with customers at the head of the queue, those customers who arrive later will remain unattended and move away. Similarly, if the subject has been primed by the experimental instructions (see procedure) to concentrate on the first few items in the memory string, later items will be registered but lost before they can be located for recall subsequently. Our interpretation of these data follows this reasoning. Comparison between the recall curves in Figs 2 and 3 for the condition with a final item of *low meaning plus suffix* may be explained accordingly. When no strict order for recalling items is stipulated in the instructions, focal attention appears to shift towards the final items at the expense of the first items in the string (Fig. 2). This shift in attention increases the probability of rehearsal and maintenance for the last item rated low in meaningfulness. Such rehearsal is effectively prevented in Expt. II where a left-to-right order of recall directs priority in recall to the first items in the string. In this case, virtually no information is available for recall for an item rated low on meaningfulness (Fig. 3). Thus, the probability of accurate recall for a registered item is increased by its continued maintenance in focal attention, as well as by an increase in the rated meaningfulness of the item itself.

Other points of similarity with visual information processing may also be cited. Turvey (1973) proposes that the process of encoding is interrupted by the onset of a patterned masking stimulus, and Jacobson (1974) demonstrated that the masking efficacy of a second information input is dependent upon its demand characteristics. This relates to our first proposal which discusses the non-interference of a non-speech suffix. Other work by Jacobson (1973) and by Coltheart & Grieve (1973) may be viewed in the context of our second proposal of registration in permanent memory. These studies demonstrated that with visual presentation and backward masking, the retrieval of the previous item is facilitated when the two items are semantically related.

In conclusion, we suggest that coding to a semantic level occurs automatically, and that this registration may act as a powerful retrieval cue. A high probability of retrieval is observed when the item is registered semantically despite the presence of a stimulus suffix, which terminates any further processing of the veridical data information.

The authors wish to acknowledge with thanks the helpful comments of David Routh on an earlier version of the paper. The authors wish to thank Miss Valerie Glendinning for her care in typing the manuscript, and to thank Mr K. Coghill and Mr G. Thompson for their technical assistance.

REFERENCES

COLE, R. A. (1973). Different memory functions for consonants and vowels. *Cog. Psychol.* **4**, 39–54.

COLE, R. A., COLTHEART, M. & ALLARD, F. (1974). Memory of a speaker's voice: reaction time to same or different-voiced letters. *Q. Jl exp. Psychol.* **26**, 1–7.

COLTHEART, M. (1972). Visual information processing. In P. C. Dodwell (ed.), *New Horizons in Psychology*, 2. Harmondsworth: Penguin.

COLTHEART, M. & GRIEVE, L. (1973). The organization of word recognition. Paper given at the British Psychological Society Meeting, December 1, 1973.

CRAIK, F. I. M. (1971). Primary memory. *Brit. med. Bull.* **27**, 232–236.

CRAIK, F. I. M. & LOCKHART, B. S. (1972). Levels of processing: a framework for memory search. *J. verb. Learn. verb. Behav.* **11**, 671–684.

CROWDER, R. G. (1969). Improved recall for digits with delayed recall cues. *J. exp. Psychol.* **82**, 258–262.

CROWDER, R. G. (1970). The role of one's own voice in immediate memory. *Cog. Psychol.* **1**, 157–178.

CROWDER, R. G. (1971). Waiting for the stimulus suffix: decay, delay, rhythm, and readout in immediate memory. *Q. Jl exp. Psychol.* **23**, 324–340.

CROWDER, R. G. (1972). Visual and auditory memory. In J. F. Kavanagh & I. G. Mattingly (eds), *Language by Ear and Eye*. Cambridge, Mass.: MIT Press.

CROWDER, R. G. & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Percept. Psychophys.* **5**, 365–373.

CROWDER, R. G. & RAEBURN, V. P. (1970). The stimulus suffix effect with reversed speech. *J. verb. Learn. verb. Behav.* **9**, 342–345.

DARWIN, C. J. & BADDELEY, A. D. (1974). Acoustic memory and the perception of speech. *Cog. Psychol.* **6**, 41–60.

DARWIN, C. J., TURVEY, M. T. & CROWDER, R. G. (1972). An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cog. Psychol.* **3**, 255–267.

JACOBSON, J. Z. (1973). Effects of association upon masking and reading latency. *Can. J. Psychol.* **27**, 58–69.

JACOBSON, J. Z. (1974). Interaction of similarity to words of visual masks and targets. *J. exp. Psychol.* **102**, 431–434.

KOLERS, P. A. (1968). Some psychological aspects of pattern recognition. In P. A. Kolers & M. Eden (eds), *Recognizing Patterns*. Cambridge, Mass.: MIT Press.

LIBERMAN, A. M., MATTINGLY, I. G. & TURVEY, M. T. (1972). Language codes and memory codes. In A. W. Melton & E. Martin (eds), *Coding Processes in Human Memory*. New York: Wiley.

LOCASCIO, D. & LEY, R. (1972). Scale-rated meaningfulness of 319 CVCVC words and paralogs previously assessed for associative reaction time. *J. verb. Learn. verb. Behav.* **11**, 243–250.

MASSARO, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychol. Rev.* **79**, 124–145.

MORTON, J. (1968). Grammar and computation in language behavior. In J. C. Catford (ed.), *Studies in Language and Language Behavior*. Center for Research in Language and Language Behavior, Progress Report No. VI. Ann Arbor, Michigan: University of Michigan.

MORTON, J. (1970). A functional model for memory. In D. A. Norman (ed.), *Models of Human Memory*. New York: Academic Press.

MORTON, J., CROWDER, R. G. & PRUSSIN, H. A. (1971). Experiments with the stimulus suffix effect. *J. exp. Psychol. Monogr.* **91**, 169–190.

NOBLE, C. E., SUTKER, P. B. & JONES, H. R. (1968). On the sigmoidal law relating association value (a) and scaled meaningfulness (m'). *Percept. mot. Skills* **26**, 375–386.

PETERSON, L. R. (1969). Concurrent verbal activity. *Psychol Rev.* **76**, 376–386.

ROUTH, D. A. (1971). Independence of the modality effect and amount of silent rehearsal in immediate serial recall. *J. verb. Learn. verb. Behav.* **10**, 213–218.

ROUTH, D. A. & LIFSCHUTZ, A. J. (1975). An asymmetrical effect of similarity in the attenuation of stimulus suffix interference. *J. verb. Learn. verb. Behav.* **14**, 95–104.

ROUTH, D. A. & MAYES, J. T. (1974). On consolidation and the potency of delayed stimulus suffixes. *Q. Jl exp. Psychol.* **26**, 472–479.

SALTER, D. (1973). Shadowing at one and at two ears. *Q. Jl exp. Psychol.* **25**, 549–556.

SALTER, D. (1975). Maintaining recency despite a stimulus suffix. *Q. Jl exp. Psychol.* **27** (In press).

SMALLWOOD, R. A. & TROMATER, L. J. (1971). Acoustic interference with redundant elements. *Psychon. Sci.* **22**, 354–356.

TURVEY, M. T. (1973). On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychol. Rev.* **80**, 1–52.

UNDERWOOD, B. J. & SCHULZ, R. W. (1960). *Meaningfulness and Verbal Learning*. Chicago: Lippincott.

APPENDIX

High and low meaningful items used in the final serial position adjacent to the suffix. The figures in parentheses refer to scale-rated meaningfulness.

High		Low		High		Low	
MONEY	(4·87)	GEMOT	(0·60)	MAN	(5·96)	QIM	(1·73)
WATER	(4·75)	WITAN	(0·60)	BED	(5·57)	ZID	(1·59)
WOMEN	(4·75)	ZIROL	(0·60)	LEG	(5·44)	JOZ	(1·58)
PAPER	(4·65)	NEROL	(0·52)	DOG	(5·43)	FOV	(1·46)
TODAY	(4·65)	VUTAW	(0·36)	WIN	(5·43)	ZIS	(1·31)
NOVEL	(4·60)	WIKOV	(0·36)	PET	(5·42)	WEF	(1·30)
RIVER	(4·60)	YOLIF	(0·36)	NUT	(5·34)	LAJ	(1·19)
JEWEL	(4·51)	ZARAO	(0·00)	TOY	(5·28)	YAV	(1·19)

From Locascio & Ley (1972). From Noble, Sutker & Jones (1968).

(*Manuscript received* 15 *December* 1974; *revised manuscript received* 27 *May* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 353–358
Printed in Great Britain

CONCURRENT CONDITIONING OF EVALUATIVE MEANING AND IMAGERY

By JEFFREY M. LOHR
Department of Psychology, University of Arkansas, Fayetteville, Arkansas 72701

Research indicates that evaluative word meaning and denotative meaning (imagery) are distinct mediational mechanisms and can be conditioned. It is hypothesized that the meaning-responses can be conditioned concurrently and independently. The present study factorially varied the unconditioned stimuli for each conditioned mediational response. Results provide support for the hypotheses and additional knowledge concerning the nature of both learning mechanisms.

The functional role of rated word imagery in learning has been most extensively investigated by Paivio (1969, 1970, 1972). It has been shown that rated ease of imagery is an important predictor of learning in a large variety of verbal learning tasks (Paivio, 1971). Word imagery is conceptualized as an associative mediational process, distinct from verbal mechanisms, which facilitates learning (Paivio, 1971).

While imagery function has been emphasized, additional research has investigated the mechanisms and variables affecting the acquisition of word imagery. Early research (Staats, Staats & Heard, 1959; Mowrer, 1960; Sheffield, 1961) conceptualized imagery as a mediational sensory-response acquired through classical conditioning. Staats (1967, 1968) has interpreted some of the findings of sensory preconditioning studies (Brogden, 1947; Bitterman, Reed & Kubala, 1953) and semantic generalization (Phillips, 1958) within the classical conditioning of imagery analysis. Other experiments have suggested the conditioning of imagery (denotative meaning). Staats, Staats & Heard (1961) demonstrated that nonsense syllables systematically paired with different words having roundness or squareness denotative meaning were differentially rated as function of the pairings. Philipchalk & Paivio (1971) showed that nonsense syllables paired with high imagery nouns were learned faster than those paired with low imagery nouns in a paired-associate learning task. The nonsense syllables were presumed to have acquired word imagery which mediated learning.

Another important word attribute, evaluative meaning, is similar in nature and acquired on the basis of the same conditioning process. Staats (1968; Staats & Staats, 1963) conceptualized evaluative word meaning as an implicit emotional meaning-response serving a mediational function. First-order classical conditioning of evaluative meaning has been shown (Staats, Staats & Crawford, 1962; Brotsky, 1968; Zanna, Kiesler & Pilkonis, 1970). Evaluative meaning can also be transferred from word to word in a higher order language conditioning paradigm (Staats & Staats, 1958; DiVesta & Stover, 1962; Paivio, 1964; Lohr & Staats, 1973; Staats, Gross, Guay & Carlson, 1973).

While word meaning and word imagery have similar mediational functions and appear to be acquired on the basis of the same conditioning process, the learning mechanisms appear distinct. The denotative meaning-response is sensory in nature (Staats, 1959, 1967, 1968) while the evaluative meaning-response is emotional in nature (Osgood, 1953, 1968; Staats, 1961, 1968). The distinction between the two mechanisms implies the meaning-responses can be acquired independently. Additionally, if the learning conditions for both evaluative meaning and denotative meaning are present, the mediational responses can be acquired concurrently.

The discussion of the relationship of word meaning and word imagery has other theoretical implications. The classical conditioning of meaning formulation has been variously criticized. Bousfield (1961) proposed that word meaning is given by the type of verbal associate elicited, making the mediational response concept unnecessary. Jenkins & Russell (1956) suggested the two different word attributes referred to the same process, corroborating Bousfield's view. However, a complex S–R analysis of the learning mechanisms (Staats, 1959; Staats, Staats, Finley & Heard, 1963) showed the correlation between word ratings and word associates was the result of simultaneous but independent conditioning processes. A similar criticism can be made on the basis of word imagery, however. Rather than being a function of a conditioned emotional response, ratings may be mediated by UCS word imagery. The findings of Philipchalk & Paivio (1971) suggest that word imagery can be transferred to previously neutral stimuli.

The purpose of the present study is to investigate the acquisition and interaction of evaluative (emotional) word meaning and denotative meaning (imagery). The present study will employ the standard language conditioning of evaluative meaning and the initial application of the first-order conditioning of denotative meaning.

It is predicted that evaluative and denotative meaning can be acquired concurrently. Additionally, it is predicted that conditioned meaning will occur independent of conditioned imagery or rated imagery of the UCS words.

METHOD

Subjects

Subjects were 74 female undergraduates enrolled in psychology courses at the University of Hawaii. Each subject received credit in her psychology class for approximately one hour's participation.

Materials

UCS words were taken from words rated on evaluative meaning (Jenkins, Russell & Suci, 1958). Words with extreme colorful/colorless ratings were eliminated. The 19 most positive, 19 most negative and 49 most neutral words were selected. CVC trigrams with low association value served as the UCS.

Conditioning lists with 72 randomly ordered CS–UCS pairs were constructed. Evaluative meaning was varied in the following way. In the first list, the CS trigram was paired with 16 positive words, one negative word and one neutral word. Three other low *a* trigrams were each paired with 16 neutral words, one positive word and one negative word. In the second list, the CS trigram was paired with 16 negative words, one positive word and one neutral word. The words of opposite evaluative meaning and the words of neutral meaning were paired with the CS trigram in each list to obscure awareness of the CS–UCS word class relationship. It is important to note that words paired with trigrams were matched on rated imagery (Paivio, unpublished) within the two lists.

Denotative meaning was manipulated by presenting different colored CS trigrams. One type of construction consisted of black line-drawn letters with white interiors on a white background. The other construction consisted of black letters with red interiors on a white background. In the experimental group (Group 1), the trigram paired with evaluative words (Trigram 1) and one of the trigrams paired with neutral words (Trigram 2) were presented primarily (i.e. 16 of 18 trials) with red interiors. The other two trigrams paired with neutral words (3 and 4) were presented primarily with white interiors. In the control group (Group 2), Trigrams 1 and 2 were presented primarily with white interiors while Trigrams 3 and 4 were presented primarily with red interiors. In this way, color exposure was equated in all groups.

The trigram factor was manipulated by constructing four additional conditioning lists in which the positions of Trigrams 1, 2 and 3 were systematically varied in each list.

Procedure

Subjects were run individually in a small experimental room. Directions were presented in written and prerecorded form. The CS–UCS relationships and the purpose of the study were obscured by an initial cover story presenting the experiment as a study of 'visual and auditory and initial learning'.

CS trigrams were projected through a one-way vision screen via a Kodak carousel projector. UCS words were recorded on a stereo tape recorder and presented via headphones. Trigram-word presentation was controlled with the use of a Kodak Sound Synchronizer. CS trigrams were exposed for 5 sec each. UCS words were presented less than 1 sec after CS onset. The interval between successive CS–UCS pairing was 0·75 sec. Subjects were given 72 exposures, 18 of each trigram.

Following the 72 trials, a sham task was given to obscure the purpose of the experiment. Subjects were directed to a table upon which rested a box with a manipulandum protruding from the top. It was explained that they were to participate in a decision-making task which 'tested' some of the things learned in the first part of the experiment. They would hear a series of spoken trigrams and were to decide to move the manipulandum toward or away. The series consisted of six trigrams presented three times each. Included were Trigrams 1, 2 and 3, their order counterbalanced across three blocks of three trials. The other three trigrams were not seen or heard previously. Feedback was provided following each manipulandum displacement. Half of the subjects were told that the 'correct' movement in response to Trigram 1 was toward them and half were told that the movement was away from them.

To summarize, subjects were randomly assigned to one of 24 groups. Half had the CS trigram paired with positive words and half with negative words. Half saw the CS trigram presented in red and half in white. Half were told correct lever movement in response to the CS trigram was toward and half were told the correct movement was away. Three different CS trigrams were used for three different groups of subjects.

Assessment of conditioning followed the auditory trigram presentation. A printed booklet was used to measure the two types of conditioned meaning and awareness. Trigrams 1, 2 and 3 were each printed above two different seven-point semantic differential-type scales. A pleasant–unpleasant scale assessed evaluative meaning and a colorful–colorless scale assessed denotative meaning. Orders of the trigram-scale combinations were determined from a random Latin square for trigram-scale combination and position. Following rating of the trigrams, additional sham tasks were administered. The first required identification of the CS trigrams within the long list of trigrams. The second was a free recall task requiring the subject to list the words presented during conditioning.

Four questions were used to assess contingency and demand awareness (see Lohr & Staats, 1973). Written responses were assessed by two blind, independent raters using predetermined criteria.

RESULTS

Eleven subjects were eliminated on the basis of rated awareness. Interrater reliability was 0·89. Sixty-three subjects were included in the statistical analysis.

Rating scores of conditioned evaluative and denotative meaning on Trigrams 1, 2 and 3 were analyzed separately in $2 \times 2 \times 2 \times 3$ ANOVAs with unequal *ns*. Observations on subjects were nested within two levels of Factor A (evaluative meaning), two levels of Factor B (denotative meaning), two levels of Factor C (movement feedback) and three levels of Factor D (syllable). The two latter factors were included in the analysis as nuisance variables. All factors were fixed. The ratings of the nonsense syllables on evaluative and denotative meaning were analyzed separately as previous factor analytic work (Frincke, 1968; Paivio, 1968) suggests the dimensions are independent. Additionally, the random assignment of trigram-scale orders to each subject partially counterbalanced for position and order effects.

Trigram 1

As predicted, pleasant–unpleasant ratings of Trigram 1 showed a significant main effect for conditioned evaluative meaning ($F = 7{\cdot}32$, d.f. = 1, 39, $P < 0{\cdot}01$). The mean rating of the positive group was 3·12 and 4·97 for the negative group, the ratings being in the predicted direction.

Colorful–colorless ratings of Trigram 1 showed a near significant effect for conditioned denotative meaning ($F = 4{\cdot}03$, d.f. = 1, 39, $P < 0{\cdot}06$). The mean rating of Group 1 was 2·89 and 3·89 for Group 2. The main C effect for movement feedback was also significant ($F = 5{\cdot}10$, d.f. = 1, 39, $P < 0{\cdot}05$), indicating the subjects receiving movement away feedback rated the trigram more colorful than subjects receiving movement toward feedback.

Trigrams 2 and 3

Pleasant–unpleasant ratings of Trigrams 2 and 3 showed no significant main effects for conditioned evaluative meaning. This finding is consistent with the theoretical position presently being tested as these trigrams were paired primarily with neutral words in both groups. It should be noted there was a significant main D effect for both Trigram 2 ($F = 4{\cdot}34$, d.f. = 2, 39, $P < 0{\cdot}05$) and Trigram 3 ($F = 14{\cdot}35$, d.f. = 2, 39, $P < 0{\cdot}001$). Newman–Keuls tests showed VUD was rated more positively than MEF or WIB when used as Trigram 2 ($P < 0{\cdot}05$) and more negatively when used as Trigram 3 ($P < 0{\cdot}01$).

Consistent with predictions, colorful–colorless ratings of Trigram 2 showed a significant main effect for conditioned denotative meaning ($F = 10{\cdot}19$, d.f. = 1, 39, $P < 0{\cdot}01$). The mean rating for Group 1 was 3·11 and 4·61 for Group 2. The main D effect was also significant ($F = 4{\cdot}43$, d.f. = 2, 39, $P < 0{\cdot}05$). The Newman–Keuls test showed trigram WIB rated more colorful than MEF or VUD ($P < 0{\cdot}05$).

Conditioned denotative meaning to Trigram 3 was predicted, but the ratings were expected in the direction opposite to those obtained with Trigrams 1 and 2. This is because Trigram 3 was presented in the other color. Results show the main B effect of ratings on the colorful–colorless scale was significant ($F = 5{\cdot}36$, d.f. =1, 39, $P < 0{\cdot}05$). The mean rating for Group 1 was 4·54 and 3·50 for Group 2. The differences were in the predicted opposite direction.

DISCUSSION

The results provide consistent support for the theoretical position that denotative meaning (imagery) can be acquired through first-order classical conditioning. The rated denotative meaning cannot be interpreted in terms of the Von Restorff effect (Wallace, 1965), as the magnitude of the effect is directly related to the degree of stimulus isolation. In the present study, total exposure to the two colors was equal in all conditioning groups. It is also appropriate to suggest that denotative ratings were a function of awareness, as has been suggested in the case of conditioned evaluative meaning (Staats, 1969). However, the use of intermittent CS-UCS schedules, sham tasks and assessed awareness of the trigram–color relationship suggests color ratings were not mediated by the experimental demand characteristics.

The results confirmed the hypothesis that evaluative word meaning and denotative meaning can be conditioned concurrently and independently. Evaluative ratings of Trigram 1 were unaffected by manipulation of denotative meaning. Similarly, ratings of denotative meaning were not affected by manipulation of evaluative meaning. Independence of mechanisms is also suggested by ratings of Trigram 2. This trigram was presented with the same color as Trigram 1. If the mediating mechanisms lacked independence, it might be predicted that evaluative meaning conditioned to Trigram 1 would generalize to Trigram 2 on the basis of common denotative meaning. However, only denotative conditioning was obtained.

The study provides additional knowledge of the learning mechanisms in evaluative and denotative meaning. By matching the rated imagery of UCS words in the conditioning groups, it is strongly suggested that conditioned emotional meaning cannot be accounted for by word imagery. Additionally, matching and elimination of extreme colorful–colorless words rules out the alternative explanation of conditioned denotative meaning on the basis of word imagery, rather than manipulation of the color stimuli. These various results are consistent with the position that evaluative meaning and denotative meaning are distinct mediational mechanisms (Staats, 1968; Paivio, 1971).

The finding that denotative ratings of Trigram 1 were affected by movement feedback is difficult to interpret. Movement feedback was included as a nuisance variable and the results have no clear relationship to theoretical issues under consideration. The finding that evaluative and denotative ratings of Trigrams 2 and 3 were affected by the nonsense syllable used suggests the need for pre-experimental assessment of CS attributes to demonstrate subtle but reliable experimental effects.

This study is based on a dissertation submitted in partial fulfillment of the requirements of the Doctor of Philosophy degree at the University of Hawaii under the supervision of Arthur W. Staats.

The author wishes to thank Dr Arthur W. Staats for helpful comments and criticism in the preparation of this report.

REFERENCES

BITTERMAN, M. E., REED, P. C. & KUBALA, A. L. (1953). The strength of sensory preconditioning. *J. exp. Psychol.* **46**, 178–182.

BOUSFIELD, W. A. (1961). The problem of meaning in verbal behavior. In C. N. Cofer (ed.), *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill.

BROGDEN, W. J. (1947). Sensory pre-conditioning of human subjects. *J. exp. Psychol.* **37**, 527–539.

BROTSKY, S. J. (1968). Classical conditioning of the galvanic skin response to verbal concepts. *J. exp. Psychol.* **76**, 244–253.

DIVESTA, F. J. & STOVER, D. O. (1962). The semantic mediation of evaluative meaning. *J. exp. Psychol.* **64**, 467–475.

FRINCKE, G. (1968). Word characteristics, associative-relatedness and the free-recall of nouns. *J. verb. Learn. verb. Behav.* **7**, 366–372.

JENKINS, J. J. & RUSSELL, W. A. (1956). Annual technical report: Basic studies on individual and group behavior. Contract No. N8 onr-66216 between the Office of Naval Research and the University of Minnesota.

JENKINS, J. J., RUSSELL, W. A. & SUCI, G. J. (1958). An atlas of semantic profiles for 360 words. *Am. J. Psychol.* **71**, 688–699.

LOHR, J. M. & STAATS, A. W. (1973). Attitude conditioning in Sino-Tibetan languages. *J. Person. soc. Psychol.* **26**, 196–200.

MOWRER, O. H. (1960). *Learning Theory and the Symbolic Processes.* New York: Wiley.

OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology.* New York: Oxford University Press.

OSGOOD, C. E. (1968). Toward a wedding of insufficiencies. In T. R. Dixon & D. L. Horton (eds), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, N.J.: Prentice-Hall.

PAIVIO, A. (1964). Generalization of verbally conditioned meaning from symbol to referent. *Can. J. Psychol.* **18**, 146–155.

PAIVIO, A. (1968). A factor-analytic study of word attributes and verbal learning. *J. verb. Learn. verb. Behav.* **7**, 41–49.

PAIVIO, A. (1969). Mental imagery in associative learning and memory. *Psychol. Rev.* **76**, 241–263.

PAIVIO, A. (1970). On the functional significance of imagery. *Psychol. Bull.* **73**, 385–392.

PAIVIO, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

PAIVIO, A. (1972). A theoretical analysis of the role of imagery in learning and memory. In P. W. Sheehan (ed.), *The Function and Nature of Imagery.* New York: Academic Press.

PAIVIO, A. Attributes of 2448 words, unpublished.

PHILIPCHALK, R. & PAIVIO, A. (1971). Acquired stimulus imagery and associative learning. *Res. Bull.* No. 183. Department of Psychology, University of Western Ontario.

PHILLIPS, L. W. (1958). Mediated verbal similarity as a determinant of the generalization of a conditioned GSR. *J. exp. Psychol.* **55**, 56–62.

SHEFFIELD, F. D. (1961). Theoretical considerations in the learning of complex sequential tasks from demonstration and practice. In A. A. Lumsdaine (ed.), *Student Response in Programmed Instruction.* (NAS-NRS Publication No. 943). Washington, D.C.: National Academy of Sciences–National Research Council.

STAATS, A. W. (1959). Meaning and word associations: Separate processes. Technical Report No. 12, Contract Onr 2305 (02) between the Office of Naval Research and Arizona State University.

STAATS, A. W. (1961). Verbal habit-families, concepts and the operant conditioning of word classes. *Psychol. Rev.* **68**, 190–204.

STAATS, A. W. (1967). Emotions and images in language: A learning analysis of their acquisition and function. In K. Salzinger & S. Salzinger (eds), *Research in Verbal Behavior.* New York: Academic Press.

STAATS, A. W. (1968). *Learning, Language and Cognition.* New York: Holt, Rinehart & Winston.

STAATS, A. W. (1969). Experimental demand characteristics and the classical conditioning of attitudes. *J. Person. soc. Psychol.* **14**, 321–328.

STAATS, A. W., GROSS, M. C., GUAY, P. F. & CARLSON, C. G. (1973). Personality and social systems and A-R-D theory: Interest (attitude) formation, function and measurement. *J. Person. soc. Psychol.* **26**, 251–261.

STAATS, A. W. & STAATS, C. K. (1958). Attitudes established by classical conditioning. *J. abnorm. soc. Psychol.* **47**, 37–40.

STAATS, A. W. & STAATS, C. K. (1963). *Complex Human Behavior.* New York: Holt, Rinehart & Winston.

STAATS, A. W., STAATS, C. K. & CRAWFORD, H. L. (1962). First order conditioning of meaning and the parallel conditioning of a GSR. *J. gen. Psychol.* **67**, 159–167.

STAATS, A. W., STAATS, C. K., FINLEY, J. R. & HEARD, W. G. (1963). Independent manipulation of meaning and *m*. *J. gen. Psychol.* **69**, 253–260.

STAATS, A. W., STAATS, C. K. & HEARD, W. G. (1959). Denotative meaning established by classical conditioning. Technical Report No. 13, Contract Nonr-2794 (02) between the Office of Naval Research and Arizona State University.

STAATS, A. W., STAATS, C. K. & HEARD, W. G. (1961). Denotative meaning established by classical conditioning. *J. exp. Psychol.* **61**, 300–303.

WALLACE, W. P. (1965). Review of the historical, empirical and theoretical status of the Von Restorff phenomenon. *Psychol. Bull.* **63**, 410–424.

ZANNA, M. P., KIESLER, C. A. & PILKONIS, P. A. (1970). Positive and negative attitude affect established by classical conditioning. *J. Person. soc. Psychol.* **14**, 321–328.

(*Manuscript received* 13 *December* 1974; *revised manuscript received* 3 *July* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 359–365
Printed in Great Britain

A FAILURE TO EXPLAIN THE EFFECTS OF FALSE HEART-RATE FEEDBACK ON AFFECT BY INDUCED CHANGES IN PHYSIOLOGICAL RESPONSE

BY E. W. THORNTON AND PATRICIA J. HAGAN
Department of Psychology, University of Liverpool

An experiment was performed to validate and extend the self-persuasion hypothesis of Valins by using slide stimuli reported to create strong unpleasant feelings. The hypothesis suggests that in situations where there is no immediate explanation of induced arousal in terms of a cue of false heart-rate feedback then reported affect is dependent on a cognitive interpretation. Evidence is presented in the present study which not only supports such an interpretation but finds no support for a previous study that reported changes in affect to be a consequence of induced changes in actual physiological responses. The implication of this discrepancy from the previous reported findings are related to the possible efficiency of techniques of systematic desensitization.

Experiments by Schachter and colleagues (Schachter, 1964; Schachter & Singer, 1962) have indicated not only that affective states are more manifest during increased sympathetic activation, but also that a cognitive interpretation of such internal sympathetic events determined the reported affective state. Despite criticisms of the experiments (Plutchik & Ax, 1967) they have generated a number of surprising studies. Valins (1966) suggested that if an individual 'thinks' he is aroused in some way then his subsequent emotionality might be based on this interpretation. Thus, if a non-veridical cognitive cue to internal state is provided for an individual his affective state should be labelled on the basis of that information. Nude photographs were presented to subjects who believed that their heart-rate was being simultaneously measured, and that a clicking sound they could hear represented their own heart-beats: in reality the clicks were presented at a false rate. When later asked to rate the nudes in terms of attractiveness the subjects, as anticipated, rated highest those slides accompanied by the clicks which changed most markedly.

Later experiments by Goldstein, Fink & Mettee (1972) and Barefoot & Straub (1971) have confirmed the essence of Valins' findings though both studies again used nude stimuli. The first replication study, however, questioned the validity of Valins' conclusions: since no evidence had been presented as to the subjects' *actual* heart-rate, the reported affect could have been induced by a real change in rate as a consequence of 'arousal' to the false-rate cue. If this were the case then the cue would not be non-veridical and the results could be interpreted within a James–Lange framework. The actual results of the Goldstein study do not clearly resolve the issue, for although the observed heart-rate changes correspond to slides accompanied by feedback changes, the ratings of attractiveness were more closely related to the false rather than the true heart-rate. An interesting, undocumented comment by Valins (1970) indicated that no evidence of an induced change was evident in his original experimental data.

The second part of the Goldstein study comprised an additional offensive condition where slides of male nudes were included in order to provide a more arousing emotion.

In this part of their experiment the ratings of how offensive the whole *proceedings* had been corresponded to the actual heart-rate changes, and the differences in false changes heard by subjects had no effect on such ratings. In this offensive condition, however, there was no attempt to relate actual heart-rate changes to the ratings of the individual slides accompanied by false feedback. The authors conclude that 'in situations not emotionally provocative a Valins effect may occur with any heart-rate mimic being an incidental by-product . . . However, in situations that are emotionally provocative, no mimic effect is observed, no Valins effect is observed and a direct relationship between actual heart-rate and emotionality is observed.' These conclusions do not invalidate Valins' observations for it is explicitly stated in Schachter's original proposals that the artificially induced arousal should have no immediate explanation. The experimental procedure of Valins and its replications derive from the premise that subjects have no simple basis for discrimination between slides other than the non-veridical cue. However, in the second part of the Goldstein study, the authors attempt to relate the non-veridical cue to a report of offensiveness in a situation where a *true* basis for such a report existed in terms of the slide material: that is, the inclusion of male nudes with female nudes. This does not mean that their conclusions are invalid either, but rather that the second part of this experiment does not permit a direct test of Valins' conclusions. The present experiment was an attempt to determine whether changes in reported affect were dependent on induced effects in heart-rate which may possibly have been absent in the relatively non-emotive situation used by Valins.

METHOD

Overview

Individuals were recruited from an experiment described as 'Heart-rate responses to unpleasant slides'. When the subject arrived at the laboratory he was told his heart-rate would be recorded while viewing ten slides of skin diseases. Two groups of subjects were led to believe that the noises they would hear emitted by the polygraph corresponded to their own heart-rate; in truth these were tape recordings. For one group (HR increase) subjects heard a systematic increase in rate during presentation of four of the slides. Subjects in the second group (HR steady) heard a stable rate throughout the recording period. A third group of subjects (Control) were shown the slides in the absence of feedback. The series of slides were shown twice with subjects actively rating them on a scale of unpleasantness during the second presentation. The actual heart-rate of each subject was recorded throughout the experimental period.

Presentation materials

The tape recordings for both false feedback groups were made from impulses emitted by a conventional physiological stimulator (Grass Instrument Co.). The pulse rate for the HR steady group tape varied every 5 sec between 66 and 72 beats per minute (BPM) throughout the experimental period. A similar baseline rate was recorded for the HR increase tape. However, during the periods when four of the slides were to be presented there was a sequential change in rate every 5 sec giving increases from 72 to 78, 84 and 90 BPM (beats/min). This was subsequently decreased to 84, 78 and 72 BPM and continued at the varying basline rate. Both tape recordings were sufficiently long to continue throughout the period of recording the true heart-rate so not to arouse any suspicion of the subject.

Ten colour photographs of skin diseases (Fry, 1973) were used as stimulus slides. The ten were selected from a larger sample on the basis of eliciting consistent feelings of 'unpleasant' and 'very unpleasant' in the largest number of subjects in a small pilot study. The slides were projected for 15 sec at 45 sec intervals. The slides were presented in synchrony with the tape recordings so that for the HR increase group the increases in rate commenced 2 sec after presentation of the third slide. Similar increases in slide–sound contingencies occurred for slides 6, 7 and 9. The positions of these 'reinforced' slides had been randomly selected from positions 3 to 10, the heart-rate being steady during the first two slide presentations to allow the subject to habituate to the procedure. Equivalent positions of slides for subjects in the two control groups are also referred to as 'reinforced' slides even though they were accompanied by either absence of feedback or a steady feedback. All remaining slide positions for each group are termed 'non-reinforced'. The second presentation of slides followed an identical pattern to the first for each subject. In order to prevent the results reflecting the intrinsic features of particular slides they were systematically rotated within all group conditions so that each slide appeared in a different position and thus each of the ten subjects in each group viewed a different sequence. The coordination of sounds and slides was automatically determined to ensure a standardized presentation to each subject within a group.

Subjects

Thirty subjects aged 19–27 years were recruited, ten subjects being randomly allocated to each of the three groups. The prerequisite of naivety and non-familiarity was assisted by non-selection of volunteers with training in psychology or medicine.

Procedure

On entering the experimental room each subject was put at ease and, after repeating the supposed purpose of the experiment, the apparatus and procedure for recording heart-rate explained. After attaching electrodes and obtaining adequate recordings the subject was given standard instructions: 'you will be shown ten slides of skin infections and the first time around I just want you to look at them. Each one will be presented for 15 sec and there will be a 30 sec interval between them. You will then be given a second showing during which you should rate each slide by assigning to it a value between 0 and 100 where 0 means that you feel it had no effect on you and 100 means you felt extremely unpleasant. The rating scale pinned in front of you will remind you of the scale to use. The rating procedure will be explained to you again during the interval between presentations, but bear in mind during the first procedure that you will be required to assign a value to each slide the second time around.'

In addition the two feedback groups were told: 'the noise you will hear is a machine that picks up each heart beat so that I can work out your average heart-rate more easily, I hope it will not be too distracting: try to concentrate on viewing the slides'.

The interval between the two presentations was three minutes during which the subjects were reminded of the task of rating each slide.

Slide ratings

The rating scale required the subject to assign a value between 0 and 100 to each slide. The scale was divided into five equal intervals ranging through neutral, mildly unpleasant, very unpleasant and extremely unpleasant. It was explained to each subject that these categories were intended only as a guideline and that a numerical value must be given to each slide. To ensure minimal disruption the subject was asked to give his ratings verbally after each slide presentation.

Heart-rate measures

The heart-rate of each subject was recorded throughout the experiment and the mean and variance in rate determined separately for each 15 sec slide presentation and the immediate preceding 15 sec. Absolute difference between the means and variance of these pre-slide and slide intervals was used as a measure of any change in rate during each presentation. These two difference scores were chosen as giving maximum information of a response change for it has been substantiated by Lacey (1967) that different people in the same situation may show great differences in the type of response change to reported and other measures of arousal. Moreover, such a difference score is preferable to comparison of rates during each slide with an overall mean rate since heart-rate could fluctuate a great deal over the experimental period.

Debriefing: adequacy of the experimental manipulation

In order to be effective as a test of the Valins phenomenon, the differential heart-rate feedback had to be accepted by the subject. The possible importance of determining the success of deception in the Valins situation has been pointed out by Golding & Lichtenstein (1970). To assess the deception the subject was questioned after the experiment. A primary, indirect series of questions was put to the subject and included such as: 'Did you find you were distracted by your heart-beats during the experiment?' Following a number of similar questions they were directly asked whether they were in doubt at any time as to the noises not being their own heart-rate. Finally, the purpose of the experiment was explained to each subject with emphasis placed on the importance of the deception. The subject was strongly urged to comment on the experiment and express doubts about the adequacy of the deception.

RESULTS AND DISCUSSION

Although the results of Golding & Lichtenstein (1970) suggest that the level of success of deception in a Valins' type experiment may not affect the results, it is clear from Schachter's propositions that the validity of results from subjects aware of a deception should be treated with suspicion and ignored. On the basis of the debriefing interviews, data from two subjects in the HR increase and the HR steady groups were omitted from the analyses. This rejection of the four subjects was performed blindly, and solely on the basis of interviews without knowledge of the individual's rating or heart-rate data. All four were confident that the sounds they heard were not their own heart-beats and voiced their suspicions during early questioning. The remaining subjects were apparently deceived and showed signs of surprise when informed of the real situation. To minimize any effects of novelty, all results to the first two slides were omitted from subsequent analyses of the data.

Slide ratings

The mean slide ratings of each group is given in Table 1. The results are in support of Valins' hypothesis. For the HR increase group the slides accompanied by heart-rate increases were rated relatively more unpleasant than those paired with a stable rate. No difference in ratings were apparent for equivalent slides in the two control groups. Non-parametric statistical analyses of these differences between groups using Kruskal–Wallis one-way analyses of variance for reinforced and non-reinforced slide positions showed a significant difference for only the non-reinforced positions ($H = 18{\cdot}25$, d.f. $= 2$, $P < 0{\cdot}001$). Subsequent Mann–Whitney comparisons were made between groups for the non-reinforced positions and also between slide positions for the three treatment groups. Adopting stringent criteria for the large number of comparisons, a significant difference was found between reinforced and non-reinforced slides for the HR increase group ($z = 2{\cdot}74$, $P < 0{\cdot}003$) and between the non-reinforced positions in the HR increase/HR steady ($z = 4{\cdot}38$, $P < 0{\cdot}00003$) and HR increase/control ($z = 4{\cdot}44$, $P < 0{\cdot}00003$) groups. All other comparisons did not *approach* significance.

Heart-rate measures

Heart-rate scores are given for only the first presentation sequence of slides when any primary effects were more likely to be unaffected by the overt verbal rating procedures.

A confirmation of the Valins effect is enhanced by an absence of a difference in heart-rate responses to non-reinforced and reinforced slide positions. A perusal of the scores revealed no consistent change in the direction of heart-rate responses to reinforced and non-reinforced slides and analysis of mean heart-rates to the two

Table 1. *The mean side ratings for non-reinforced and reinforced slide positions for each group of subjects*

	HR increase	HR steady	Control
Reinforced	44·59	47·88	54·25
Non-reinforced	31·34	46·66	55·27
Difference	13·25	1·22	1·02

Table 2. *The mean difference heart-rate scores between pre-slide and slide periods for the two difference scores for each group of subjects during reinforced and non-reinforced slide positions*

		HR increase	HR steady	Control
Reinforced	Mean difference score	3·047	1·872	2·708
	Variance difference score	26·188	18·93	17·407
Non-reinforced	Mean difference score	3·609	3·203	2·175
	Variance difference score	21·131	14·897	17·142

sets of slides confirmed an absence of any significant difference in rate between reinforced and non-reinforced positions.

However, since individuals show great variability in any channel of autonomic activity in terms of magnitude, direction and variance (Lacey, 1967), it could be suggested that an absence of any differences in the magnitude of *disturbance* in heart-rate to reinforced and non-reinforced slides would provide a more stringent interpretation of Valins' work. Thus, a test of this null hypothesis was based on the absolute differences between mean and variance scores for pre-slide and slide presentations. Table 2 presents the means scores of both measures for both reinforced, and non-reinforced slide positions for the three treatment groups. An absence of homogeneity of variance between the groups precluded an overall parametric analysis and, after transforming the data for normalization, two factor analyses of variance were carried out for each group separately for both measures. The two factors were slide position and slide conditions (non-reinforced and reinforced). These analyses showed strong insignificant main effects and interactions ($F \leqslant 1$) for each group in both the difference and variance measures. Thus, there was no evidence for induced changes in rate and, subsequently, for the differential ratings of slides in terms of true heart-rate response. Further analyses of these heart-rate changes during the second presentation of slides, despite contamination by active rating, confirmed these insignificant findings.

Although the ratings of unpleasantness in studies should be treated with caution since they may reflect different dimensions of the elusive construct of emotion, and such factors as conformity to expectancies may interact or substitute for felt affect, the results of the present study contrast with the conclusions of Goldstein *et al.*, and re-establish the validity of Valins' conclusion. However, the results do suggest an alternative effect of a cognitive interpretation of false cues of arousal not overtly considered by Valins but which is central to the subsequent extensions of his hypothesis for the techniques of systematic desensitization. The interpretation could be considered a corollary of the self-persuasion hypothesis. Specifically, although the false cue enhances the relative level of unpleasantness of slides accompanied by an apparent change in rate for subjects in the experimental group, comparison of the ratings with control groups suggest that the effect is produced by a decrease in affect to those slides accompanied by steady feedback. A comparison with the rating data from the group receiving steady feedback to all slides confirms that this effect is not a consequence of steady feedback alone, but occurs only when there is an opportunity to contrast responses to slides accompanied by changes in rate. In a situation allowing such comparison, slides accompanied by cues suggesting 'no arousal' are perceived and interpreted as affectively more neutral.

It would appear that false feedback as an inferred index of arousal may be a determinant of affect in a potent emotional situation, providing the subject has no other immediate explanation for his apparent reactions. However, it may be that the differential effectiveness of attempts to manipulate affect using false cues is dependent upon the extent to which these changes are correlated with changes in actual rate, and the discriminability of the specific stimulus information with respect to such changes. Such a proposition might possibly explain discrepant results from studies which have postulated the importance of cognitions, inferred directly from false heart-rate feedback or indirectly from progress up a hierarchy of stimuli, as determinants of the efficiency of systematic desensitization. Such studies have obtained contradictory results both internally and across studies: Valins & Ray (1967) and Marcia, Rubin & Efran (1969) found evidence for cognitive manipulation of avoidance behaviour though no change in reported affect. Conversely, Wilson (1973), though presenting evidence against a cognitive interpretation, provided data suggesting that there could be such a component determining the reported affect but not behaviour. The deficiency of these studies is that there has been no measurement of actual change in physiological arousal and one cannot, therefore, determine if changes in the different dimensions of emotionality are correlated with either the false cue or actual indices of arousal. It would even be consistent with Schachter's (1971) work, if it were found that individuals are differentially susceptible to both false and true peripheral physiological responses.

REFERENCES

BAREFOOT, J. C. & STRAUB, R. B. (1971). Opportunity for information search and the effect of false heart-rate feedback. *J. Person. soc. Psychol.* **17**, 154–157.

FRY, L. (1973). *Dermatology: An Illustrated Guide*. Update Publications.

GOLDING, S. L. & LICHTENSTEIN, E. (1970). Confession of awareness and prior knowledge of deception as a function of interview set and approval motivation. *J. Person. soc. Psychol.* **14**, 213–223.

GOLDSTEIN, D., FINK, D. & METTEE, D. R. (1972). Cognition of arousal and actual arousal as determinants of emotion. *J. Person. soc. Psychol.* **21**, 41–51.

LACEY, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumball (eds), *Psychological Stress*. New York: Appleton-Century-Crofts.

MARCIA, J. E., RUBIN, B. M. & EFRAN, J. S. (1969). Systematic desensitization: Expectancy change or counterconditioning? *J. abnorm. Psychol.* **74**, 382–387.

PLUTCHIK, R. & AX, A. F. (1967). A critique of 'Determinants of emotional state' by Schachter and Singer. *Psychophysiology* **4**, 79–82.

SCHACHTER, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, vol. 1, pp. 49–80. New York: Academic Press.

Schachter, S. (1971). *Emotion, Obesity and Crime*. New York: Academic Press.

Schachter, S. & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychol. Rev.* **69**, 379–399.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *J. Person. soc. Psychol.* **4**, 400–408.

Valins, S. (1970). The perception and labelling of bodily changes as determinants of emotional behaviour. In P. Black (ed.), *Physiological Correlates of Emotion*, pp. 229–243. New York: Academic Press.

Valins, S. & Ray, A. (1967). Effects of cognitive desensitization on avoidance behaviour. *J. Person. soc. Psychol.* **7**, 345–350.

Wilson, G. T. (1973). Effects of false feedback on avoidance behaviour. 'Cognitive' desensitization revisited. *J. Person. soc. Psychol.* **28**, 115–122.

(*Manuscript received* 27 *January* 1975; *revised manuscript received* 3 *May* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 367–375
Printed in Great Britain

THE ELECTRODERMAL COMPONENT OF THE ORIENTING RESPONSE IN BLIND AND DEAF INDIVIDUALS

By D. CARROLL and P. G. SURTEES
Department of Psychology, University of Birmingham

Skin conductance response (SCR) was monitored as an index of the orienting response (OR) to a series of 20 tones presented to blind and sighted schoolchildren. Blind children showed fewer criterion SCRs than sighted children, and significantly more of the blind sample exhibited SCR habituation. The blind group also demonstrated fewer non-specific SCRs than sighted control subjects, but demonstrated a higher average skin conductance level (SCL). A second experiment found similar effects for deaf versus hearing children presented with simple visual stimulation, save that no significant difference was observed between these groups in SCL. These results can be explained by assuming that loss of modality results in more efficient development of a neuronal model of the salient features of stimuli impinging on the remaining intact modalities. An explanation in terms of the assessed significance of the stimulus situation can also be entertained. The atypical skin conductance activity of three of the blind group was given separate consideration.

When presented with a novel stimulus an individual will initially manifest an orienting response (OR) as indicated, for example, by a skin conductance response (SCR) and peripheral vasoconstriction. With repetition of the same stimulus, the OR will habituate. Sokolov (1963) proposed an association between the OR and heightened general perceptual sensitivity. He cited several studies in which an OR elicited by a tone resulted in a decrease in visual threshold. As the OR to the tone gradually habituated, visual threshold was restored to its initial level. However, the acquisition by the extraneous auditory stimulus of signal significance resulted in a marked OR and an associated rise in visual sensitivity, which persisted after many stimulus presentations. Evidence shows that the investment of signal significance in a stimulus increases the magnitude and resistance to habituation of the OR (Sokolov, 1963; Germana & Chernault, 1968; Maltzman & Mandell, 1968).

Although it is widely held that blind individuals compensate for lack of sight by increased sensitivity in the remaining sensory modalities, this claim has received little in the way of controlled experimental attention. However Riley, Luterman & Cohen (1964) reported that highly mobile blind individuals were very sensitive to auditory frequencies outside the normal speech range. Verbuk (1957) and Fishlev (1959), showed that the blind are better able to use a conditional code made up of auditory and tactile stimuli than are normal sighted individuals. These results add weight to notions concerning increased sensitivity in the remaining sensory modalities of the blind.

Measurements of the OR, having established associations with perceptual sensitivity, provide a simple, alternative means of evaluating the sensitivity of the blind to information impinging on intact modalities. Zemtsova (1969) indicated that '. . . loss of vision causes a restructuring of the interanalyzer connections and an alteration of the signal significance of many stimuli' (p. 304). If the blind do confer signal significance on stimuli impinging on the remaining intact modalities, then it would be expected that habituation of the components of the OR, elicited, for example, by auditory stimuli, would be slower in blind than in sighted individuals. Asakhov, Zimkina & Stepanov (1955), in fact, demonstrated that the blind manifest an intensified OR to simple auditory stimuli and maintain this for an extended period of time. However there are contradictions in the Soviet literature as regards the nature of the OR in blind individuals. Zemtsova, Novikova, Paramonova & Sokolov (1965) found that with simple auditory stimuli, presented through headphones, the SCR component of the OR in the blind habituated in three to five presentations, whereas in normal sighted individuals, habituation occurred after seven to nine stimulus presentations.

In view of these conflicting findings regarding OR habituation in blind individuals, further experimentation seems essential. Thus Expt. I explores habituation of the SCR component of the OR to simple auditory stimuli in blind and sighted schoolchildren. Expt. II examines the behaviour of the deaf in an analogous situation. The nature of SCR habituation to simple visual stimulation is compared for deaf and hearing children.

EXPERIMENT I

Method

Subjects

Subjects were two groups of 11 schoolboys. The experimental group was totally blind. Seven of the group had been blind from birth, two from early infancy (less than 5 months), one from the age of 2 years and one from 5 years. Apart from blindness, these children exhibited no other sensory handicap and had no recorded neurological pathology, except one boy who showed some spasticity. Three of the group were rated by the Worcester College for the Blind as showing excellent mobility and orientation, six as showing good and two as showing average. The average age of the blind children was 12 years 9 months (s = 9 months). Their average verbal IQ, given by the college, was 119·4 (s = 8·13). The control group of schoolchildren had normal sight. These children had an average age of 13 years (s = 4 months). Average IQ as measured by the WAIS was 121·4 (s = 7·4).

Apparatus and procedure

Subjects were presented with a series of 20 1000 Hz, 75 dB (re: 0·0002 dyne/cm^2) pure tones. Each tone was of 3·5 sec duration, and the inter-stimulus interval varied randomly, from 12 to 24 sec with a mean of 18 sec.

These orienting stimuli were presented by means of a punched tape and automatic tape reader. An audio generator produced the tones which were then delivered through Akai headphones. SCR was measured using a portable two channel Beckman RS type dynograph and a Beckman SC coupler. The remaining dynograph channel provided a marker of tone onset and offset.

The experiment was conducted in rooms made available by the schools concerned. Subject semi-reclined in a comfortable armchair and formal instructions were given concerning tone presentation and physiological measurement. The experimenter asked the subject to relax and to remain as still as possible throughout the experimental session. Ag/AgCl recording electrodes were then attached to the palm and second digit of subject's left hand. After impressing upon the subject that no behavioural response would be required of him the headphones were put on. A medical screen separated subject from the recording apparatus. Subject was now allowed five minutes in which to relax before stimulus presentation commenced. The whole procedure including debriefing lasted approximately 20 minutes.

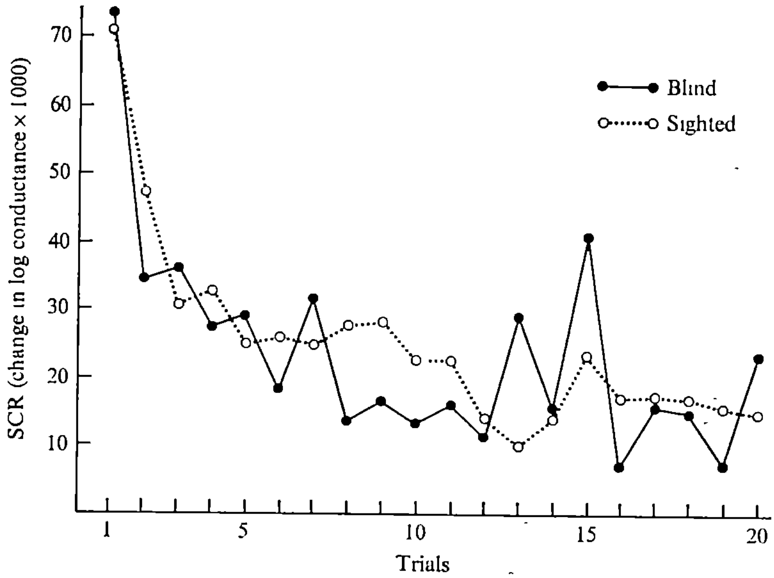


Fig. 1. Mean skin conductance response (SCR) to a brief auditory stimulus on each of 20 trials for blind and sighted schoolchildren.

Results

Skin conductance response

Decreases in skin resistance of 500 ohm or greater with a latency of 1–5 sec after stimulus onset were scored as SCRs evoked by the stimulus. The criterion is suggested by O'Gorman, Mangan & Gowen (1970). The resistance level immediately prior to tone onset and the minimum resistance level reached after tone onset were noted. These values were transformed into log conductance units and change scores subsequently computed. Lader (1964), Montagu & Coles (1966) and Glass & Singer (1972) recommend the representing of SCRs to repeatedly presented stimuli in terms of changes in log conductance.

These change scores were averaged across subjects in each group and the resultant mean values are plotted in Fig. 1. A two factor analysis of variance (repeated measures on one factor) was performed on the SCR data. Only the trials factor had a significant effect ($F = 6{\cdot}44$, d.f. = 19, 380, $P < 0{\cdot}001$).

However, inspection of the SCR data indicated that whereas the average SCR during the later trials was a representative indicator of group behaviour for the sighted control group, this average for the blind group represented the behaviour of only three blind subjects who persisted in exhibiting extremely large SCRs to the later tone presentations. Thus it was decided to analyse SCR data in terms of the number of criterion SCRs (i.e. resistance changes greater than 500 ohm) exhibited by each group of subjects to each tone presentation. The total number of responses so computed for each group on each trial are presented in Fig. 2. Inspection of Fig. 2 indicates that fewer blind than sighted control subjects exhibited criterion SCRs. A Mann–Whitney U test indicated that the distribution of criterion SCRs across trials differed significantly for the two subject groups ($U = 80{\cdot}5$, $P < 0{\cdot}002$). Employing an SCR habituation criterion of three trials without response, the number of subjects in each group showing habituation was determined (O'Gorman *et al.*

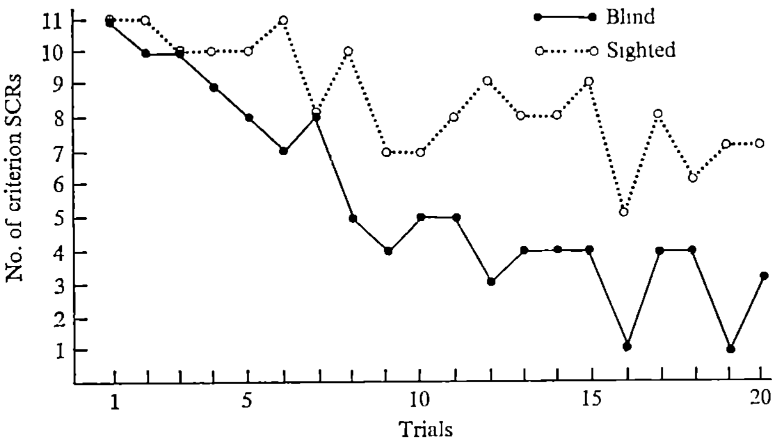


Fig. 2. Total number of criterion skin conductance responses (SCRs) shown on each trial of a brief auditory stimulus by the blind and sighted groups (criterion = resistance change greater than 500 ohm, occurring 1–5 sec after stimulus onset).

Table 1. *Average SCLs in μmhos at each of the sampling times for blind and sighted subjects*

	0·0	1·5	3·0	4·5
Blind	16·34	16·30	17·31	17·77
Sighted	9·82	9·59	10·08	10·48

1970). Whereas nine blind subjects demonstrated habituation, only three sighted subjects did so. However, for one of the habituating blind subjects, the three consecutive response failures occurred early on in the stimulus series, and, immediately following these omissions, this subject continued to exhibit large SCRs to the tone presentations. Thus, only eight blind subjects can be considered as habituators. Even so, a Fisher's exact probability test indicated that the blind and control groups differed significantly with respect to the number of subjects showing SCR habituation ($P < 0{\cdot}05$).

Non-specific electrodermal fluctuations

The number of non-specific SCRs emitted by subjects during the tone series was recorded. These non-specific fluctuations were defined as phasic decreases in skin resistance of at least 500 ohm not occurring within 1–5 sec of stimulus onset. The mean number of such non-specific SCRs was 32·00 ($s = 20{\cdot}65$) for the sighted group and 15·75 ($s = 10{\cdot}00$) for the blind group. A two-tailed t test was conducted on such values following square root transformation to reduce heterogeneity of variance, and revealed that the difference between the two groups in spontaneous SCR activity was significant ($t = 3{\cdot}23$, d.f. $= 20$, $P < 0{\cdot}01$).

Skin conductance level

Skin conductance level (SCL) was sampled for each subject at 0, 1·5, 3, 4·5 min from a starting point 10 sec prior to the delivery of the first tone in the habituation series. The average SCLs at each of these sampling points for the blind and sighted subjects are presented in Table 1. Analysis of variance indicated that SCL differed significantly for the two groups ($F = 4{\cdot}81$, d.f. = 1, 20, $P < 0{\cdot}05$). Inspection of the SCL data indicated that three blind subjects exhibited very large SCLs (average SCL for the three was 24·49 μmhos). It is noteworthy that it was these same blind subjects who demonstrated persistently large SCRs to the tones compared to their group average. There was a slight tendency for SCL to change during the tone series ($F = 2{\cdot}19$, d.f. = 3, 60, $P < 0{\cdot}10$). The interaction effect was not significant ($F = 0{\cdot}38$).

Experiment II

Method

Subjects

Subjects were two groups of 12 schoolboys. The experimental group was, for the purposes of this present study, considered as totally deaf. An analysis of audiometric tests given to each of the deaf subjects is shown in Fig. 3. Inspection of Fig. 3 indicates that what hearing the children possessed was specific to certain frequencies and that the necessary intensity of sound to produce a response at these frequencies had to be very high. For many of the subjects no response was obtained even when the intensity was raised to 110 dB over the full range of frequencies. All 12 children in this group had been established as deaf since early childhood. Apart from deafness the children exhibited no other sensory or neurological handicap. The average age of this group was 13 years 1 month ($s = 12$ months). Their average IQ, as gauged by the Snyder–Oomen test, was 103·7 ($s = 8{\cdot}3$). The control group of schoolchildren had normal hearing. Their average age was 13 years 4 months ($s = 9$ months) and their average IQ as gauged by Standard Progressive Matrices was 101·5 ($s = 9{\cdot}9$).

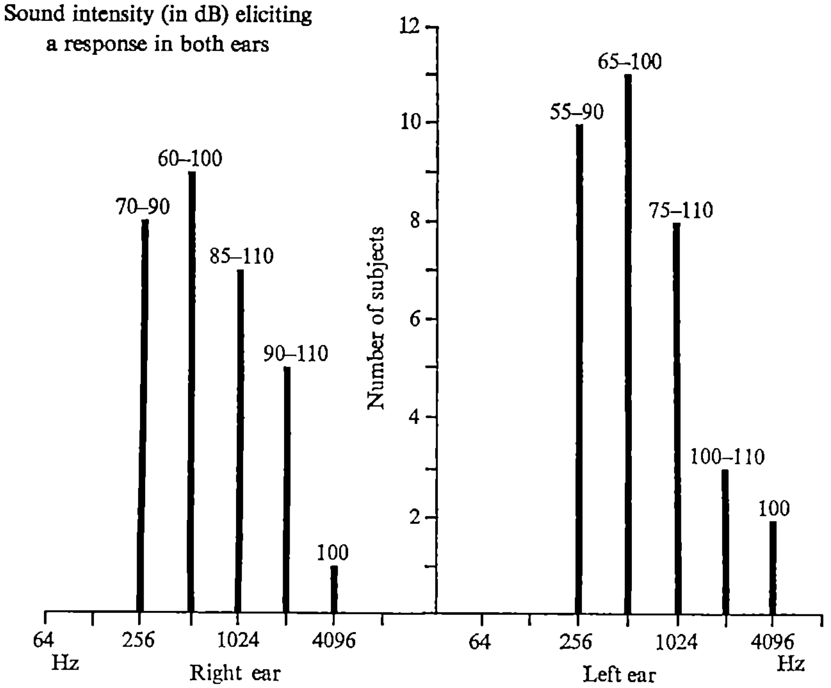


Fig. 3. Summary of audiometric analyses for deaf group.

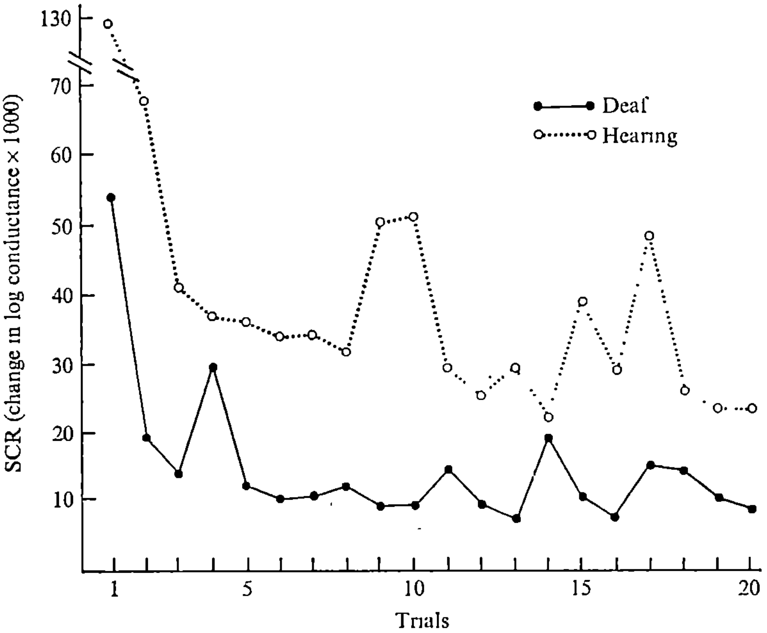


Fig. 4. Mean skin conductance response (SCR) to a brief visual stimulus on each of 20 trials for deaf and hearing schoolchildren.

Procedure

The procedure was identical to that used in Expt. I, save that subjects received visual stimulation. The visual stimulus consisted of a 3·5 sec duration beam of light directed at the subject's face. At the point where subject received the stimulus, the intensity of illumination of the source was 10 foot-candles. Instructions were similar to those used previously except that an intermediary communicated with the deaf sample. Headphones were worn by the control group to attenuate ambient noise level.

Results

Skin conductance responses

SCR was computed as before, the mean values for each group are plotted in Fig. 4. Analysis of variance revealed a significant effect of group ($F = 9{\cdot}12$, d.f. = 1, 22, $P < 0{\cdot}01$) and of trials ($F = 7{\cdot}23$, d.f. = 19, 418, $P < 0{\cdot}001$) on SCR. The groups × trials interaction was also significant ($F = 2{\cdot}13$, d.f. = 19, 418, $P < 0{\cdot}025$).

The number of criterion SCRs was noted as before. The total such SCRs on each stimulus trial for each group is presented in Fig. 5. A Mann–Whitney U test again indicated that the distribution of criterion SCRs across trials differed for the two groups ($U = 58$, $P < 0{\cdot}001$). The number of subjects in each group showing habituation to the criterion of three consecutive response failures was computed as before. Ten of the deaf sample, but only two of the hearing group, showed SCR habituation. A Fisher's exact probability test revealed that deaf and hearing groups differ with respect to the number of subjects showing SCR habituation ($P < 0{\cdot}002$).

Non-specific electrodermal fluctuations

The number of non-specific SCRs emitted by subjects during the tone series was counted. Whereas the deaf children exhibited on average 16·25 ($s = 12{\cdot}36$) such

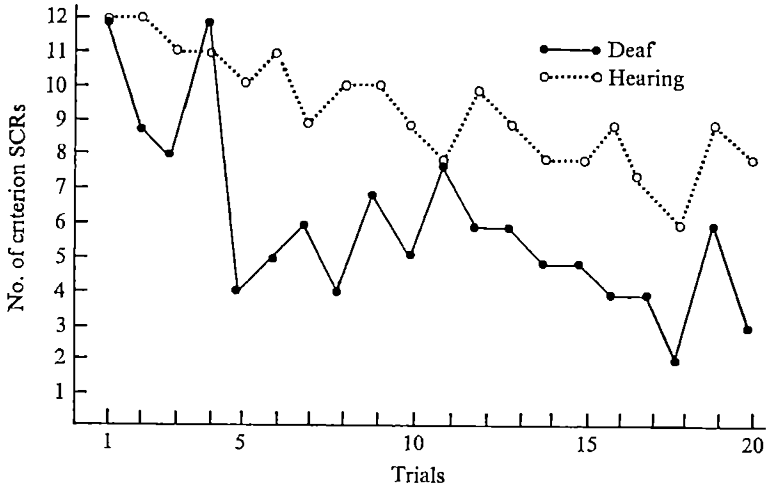


Fig. 5. Total number of criterion skin conductance responses (SCRs) shown on each trial of a brief visual stimulus by the deaf and hearing groups (criterion = resistance change greater than 500 ohm, occurring 1–5 sec after stimulus onset).

Table 2. *Average SCLs in μmhos at each of the sampling times for deaf and hearing subjects*

	0·0	1·5	3·0	4·5
Deaf	8·80	9·18	10·38	10·31
Hearing	13·14	11·50	11·80	13·37

fluctuations, the hearing group average was 42·50 (s = 25·32). Square root transformation was applied to the spontaneous SCR data as before. A t test indicated a significant difference between deaf and hearing children in spontaneous SCR activity exhibited during the stimulus series (t = 3·37, d.f. = 22, P < 0·01).

Skin conductance level

SCL was sampled as before. The average SCLs for each group are presented in Table 2. Analysis of variance was applied to the SCL data, but no significant effects were observed.

Discussion

The results of these two experiments indicate that the loss of one modality generally results in the subject's showing successively less in the way of SCR ORs to a series of simple stimuli presented to a remaining intact modality. Further, complete SCR habituation to the stimuli was a more frequent occurrence for the blind and deaf groups than their respective control samples. As such these findings are at odds with the earlier results of Asakhov *et al.* (1955), but are in line with those of Zemtsova *et al.* (1965).

Sokolov (1960) proposed that the OR occurs where there is a mismatch between stimulus input and the neuronal representation or model of the stimulus world. OR habituation results from the gradual emergence, with stimulus repetition, of a neuronal model of the salient features of stimulus input, such as intensity, duration and the interval between successive stimuli. It is possible that blind and deaf individuals, because of their greater reliance on alternative modality stimulation, are better able to discern salient stimulus features than individuals with all modalities intact. Thus the blind and deaf would develop neuronal models of simple auditory and visual stimuli (respectively) more efficiently than would normal controls. This could account for the more frequent SCR habituation observed in the present study for the blind and deaf schoolchildren. The efficient abstraction of features of auditory and tactile stimulation by blind individuals shown by Verbuk (1957) and Fishlev (1959) supports such an interpretation.

Bernstein (1973) has shown that subjects listening to auditory stimuli which have behavioural significance or signal value exhibited relatively high frequencies of spontaneous electrodermal activity compared to subjects exposed to auditory stimuli without such significance. The present observation of fewer spontaneous SCRs exhibited by blind and deaf subjects suggests that these individuals assessed the stimuli and perhaps the whole experimental situation as less significant than control subjects. Interpretations in terms of decreased significance and more efficient stimulus processing, however, are not mutually exclusive; the latter may precipitate the former.

An alternative explanation of the observed SCR effects in terms of the blind and deaf subjects being unable to sustain tonic arousal during the stimulus series seems unlikely in the context of the SCL data. No relative decline in SCL was observed for these subjects. In fact, the blind group showed higher SCLs than sighted control subjects throughout the experimental session and generally SCL, if anything, rose slightly as the session progressed. According to Malmo (1965), rising gradients for autonomic parameters are associated with the maintenance of cortical tone and attention.

While blind and deaf groups generally showed less OR activity to the stimuli than normal control subjects, three blind subjects demonstrated SC activity which diverged considerably from this pattern. These subjects showed exceptionally large SCRs to the auditory stimuli, particularly to those later in the stimulus series. Further, the high average SCL shown by the blind group was contributed to substantially by these three subjects. Gruzelier & Venables (1972) have argued that the occurrence of high SCLs and persistently large SCRs to repetitive stimulation in schizophrenic subjects indicate limbic system dysfunction. More precisely, dysfunction of the hippocampal inhibitory system is invoked. The atypical SC activity of the three blind subjects in the present study could reflect such a pathology of the orienting system. It is worth noting that the two subjects rated as only average in general mobility and orientation were included in the three.

The cooperation of the following schools, their pupils and their Head Teachers is gratefully acknowledged: the Worcester College for the Blind, Worcester; the Royal School for Deaf Children, Birmingham; Dame Elizabeth Cadbury Comprehensive, Birmingham; the Blessed Humphrey Middlemore Comprehensive, Birmingham.

Dr D. Carroll, to whom requests for reprints should be addressed, gratefully acknowledges the receipt of a Medical Research Council Junior Research Fellowship at the time of conducting this study.

REFERENCES

ASAKHOV, B. D., ZIMKINA, A. M. & STEPANOV, A. I. (1955). The peculiarities of the orienting reaction to audio-stimulation in the blind. *Fiziol. Zh. U.S.S.R.* **41**, 314–320.

BERNSTEIN, A. S. (1973). Electrodermal lability and the OR: Reply to O'Gorman and further exposition of the 'significance hypothesis'. *Aust. J. Psychol.* **25**, 147–154.

FISHLEV, Y. R. (1959). Substituting for visual functions in the education of blind children by use of photoelectric equipment. Candidate's dissertation, Sverdlosk. Cited in M. Cole & I. Maltzman (eds), (1969), *A Handbook of Contemporary Soviet Psychology*. New York: Basic Books.

GERMANA, J. & CHERNAULT, G. (1968). Patterns of galvanic skin response to signal and non-signal stimuli. *Psychophysiology* **5**, 284–292.

GLASS, D. C. & SINGER, J. E. (1972). *Urban Stress*. New York: Academic Press.

GRUZELIER, J. H. & VENABLES, P. H. (1972). Skin conductance orienting activity in a heterogeneous sample of schizophrenics. *J. nerv. ment. Dis.* **155**, 277–287.

LADER, M. H. (1964). The effect of cyclobarbitone on habituation of the psychogalvanic reflex. *Brain* **87**, 321–340.

MALMO, R. B. (1965). Physiological gradients and behavior. *Psychol. Bull.* **64**, 225–234.

MALTZMAN, I. & MANDELL, M. P. (1968). The orienting reflex as a predictor of learning and performance. *J. exp. Res. Personality* **3**, 99–106.

MONTAGU, J. D. & COLES, M. E. (1966). Mechanism and measurement of the galvanic skin response. *Psychol. Bull.* **65**, 261–279.

O'GORMAN, J. G., MANGAN, G. L. & GOWEN, J. A. (1970). Selective habituation of the galvanic skin response component of the orientation reaction to an auditory stimulus. *Psychophysiology* **6**, 716–721.

RILEY, L. H., LUTERMAN, D. M. & COHEN, M. F. (1964). Relationship between hearing ability and mobility in a blinded adult population. *New Outlook* **58**, 139–141.

SOKOLOV, E. N. (1960). Neuronal models and the orienting reflex. In M. A. Brazier (ed.), *The Central Nervous System and Behavior*. New York: Macey.

SOKOLOV, E. N. (1963). *Perception and the Conditioned Reflex*. Oxford: Pergamon Press.

VERBUK, M. A. (1957). New possibilities for increasing palpability in schools for the blind. In *Teaching and Training Work in Special Schools*. Moscow: Uchpedgiz. Cited in M. Cole & I. Maltzman (eds), (1969), *A Handbook of Contemporary Soviet Psychology*. New York: Basic Books.

ZEMTSOVA, M. I. (1969). Characteristics of perceptual activity in the blind. In M. Cole & I. Maltzman (eds), *A Handbook of Contemporary Soviet Psychology*. New York: Basic Books.

ZEMTSOVA, M. I., NOVIKOVA, L. A., PARAMONOVA, N. P. & SOKOLOV, E. N. (1965). Investigation of the orienting reflex to sound stimuli in blind persons. In L. G. Veronin, A. N. Leontiev, A. R. Luria, E. N. Sokolov & O. S. Vinogradova (eds), *Orienting Reflex and Exploratory Behavior*. Washington, D.C.: American Institute of Biological Sciences.

Manuscript received 30 *May* 1973; *revised manuscript received* 15 *May* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 377–390
Printed in Great Britain

FREQUENCY, CONCEPTUAL STRUCTURE AND PATTERN RECOGNITION

By J. G. WOLFF
University Hospital of Wales and University College, Cardiff

The frequency of co-occurrence of perceptual elements has been employed by Wolff (1975) as an explanatory principle in a model of speech segmentation. Here a computer program is described which uses the same principle to model concept formation. A second program is also described which recognizes or categorizes new patterns using the 'conceptual structure' developed by the first program. Six features of human conceptual/recognition systems are modelled with varying degrees of success: Salience of concepts; hierarchical relations amongst concepts; overlap between concepts; 'fuzziness' of conceptual boundaries; the polythetic nature of human concepts including the possibility of recognizing patterns in spite of distortions; differential weighting of attributes in recognition.

The model incorporates coding by 'schema plus correction' proposed by Attneave (1954) and Oldfield (1954) as a means of effecting economical storage of information. However the particular form of this principle used here achieves economical coding only with certain types of data. A possible means of overcoming this problem is briefly considered.

The absolute joint probability of perceptual elements (or, equivalently, their frequency of co-occurrence) has been employed by Wolff (1975) as an explanatory principle in a model of speech segmentation. In that context it had the virtue of suggesting a basis for the perceptual coherence of linguistic entities, at the same time permitting a very efficient coding of the data. The purposes of this paper are firstly to show that this same principle has explanatory value in a model of human concept formation but also to demonstrate that it can sometimes lead to serious diseconomies in coding. It seems that a related principle must be found to overcome this problem whilst preserving the advantages of the original principle.

Two computer programs are described: CLST02 which builds up a 'conceptual structure' and RKLS02 which uses that structure to recognize or categorize patterns.

Apart from the above reference the affinities of this work are threefold. Uttley (1959, 1970) has explored the relation between frequency and classification and there are similarities between his models and the one described here. Secondly, the model is related to the large new field of numerical taxonomy which is concerned with methods of classifying objects (or organisms) by their properties or attributes (see Cole, 1969; Jardine & Sibson, 1971; Sneath & Sokal, 1973). The model differs from the bulk of clustering algorithms in that it produces overlapping rather than discrete clusters. It also differs importantly from all methods which start with, or start by forming, a similarity matrix for the objects to be classified which is then the sole basis for subsequent data manipulations. This is the chief feature distinguishing it from methods developed at the Cambridge Language Research Unit (e.g. Needham, 1963; Spärck Jones & Jackson, 1967, 1970). Perhaps the model's closest relative is McQuitty's (1956) Agreement Analysis but it differs from this in several respects the most important of which is its explicit handling of the recognition problem.

The third area of relevance is 'schema theory' which embodies the supposition that much of cognition can be seen as the use of redundancies in sensory information to effect economies in the storage and retrieval of that information. Several possible methods of achieving economies have been set out by Attneave (1954, 1957) and similar ideas have been discussed by Oldfield (1954). More recently workers at Texas Christian University and elsewhere have taken up the idea of 'schema plus correction' as a means of economical coding of information (see, for example, Evans, 1964, 1967; Posner & Keele, 1968; Brown & Evans, 1970; Bersted, 1970; Evans & Ellis, 1972). Wallace & Boulton (1968) and Boulton & Wallace (1970) have related clustering to economical coding of information but not in terms of 'schema plus correction'.

The exact nature of a 'schema' and of 'corrections' to that schema remains a problem. Posner & Keele (1968) write that 'The philosophical notion of abstract ideas is vague but it does suggest that information which is common to the individual instances is abstracted and stored in some form. In its strongest sense, this might be translated operationally into the hypotheses that the commonalities among a set of patterns are abstracted during learning and that they alone are stored' (p. 354).

Given the utility of the frequency of co-occurrence principle in the speech segmentation model, CLST02 was designed to explore the implications of defining a schema as a set of commonly co-occurring attributes and corrections to a schema as the addition of other attributes which co-occur with the schema set but with a lower frequency.

TERMS AND ASSUMPTIONS

Prior to the application of the model it is assumed that the perceptual world has been segmented into discrete entities (objects, words, actions, etc.). It is not intended to discuss this segmentation except to remark that the processes described by Olivier (1968), Zahn (1971) and Wolff (1975) are possible mechanisms for achieving this partitioning. In the jargon of numerical taxonomy these entities are termed operational taxonomic units (OTUs) but the term 'entity' or, more simply, 'object' will serve our purpose.

The second assumption is that each object has a set of discrete properties or attributes. Here, for simplicity, it will be assumed that each attribute has only two states, 'present' or 'absent'. Nothing in the model precludes an hierarchical relation between attributes: both 'hand' and 'fingers' may be given as attributes even though the latter may presuppose the former.

A third assumption, common to most clustering algorithms, is that it is meaningful to treat the statistical interrelations of attributes independently of their spatial or temporal interrelations. The lists of attributes for each object are unordered lists although it is clear that the order or arrangement of attributes of real objects has a bearing on categorization or recognition. This then is a simplifying assumption made for want of a satisfactory solution to this problem.

Before proceeding to describe the model it is necessary to define the notions of concept and recognition and to specify those properties of human conceptual/recognition systems which the program is intended to model. For the present purpose the term 'concept' which is here used interchangeably with 'category' or 'class' will designate a set of entities the members of which are in some sense similar; what that sense is should become apparent from the description of the model. The recognition problem is the problem of assessing a new previously unobserved object and deciding to which conceptual set or sets it should properly be assigned even though it may not be identical with any previously observed object. Of the six properties of a conceptual system to be described here the last three are most directly related to recognition.

Salience. In common with Brown & Evans (1970) and Rosch (1973) it is supposed that certain categories of object or other perceptual entity are widely or universally employed by people (and perhaps animals too – see Premack, 1970) because they are salient or natural groupings. Of the many possible ways of classifying people, say, such divisions as male/female are employed in preference to arbitrary or bizarre divisions such as 'people who have fair hair and a handspan greater than seven inches/those who do not'. The salience of a category is assumed to be related in some way to the intercorrelation among the attributes which determine the category.

The notion of salience and the notion that concepts may vary in salience implies that concepts are developed by a process of sifting and sorting of sensory material carried out autonomously. This view of concept acquisition should be distinguished sharply from the notion of 'concept attainment' in which subjects discover arbitrary concepts with the help of a 'teacher' (e.g. Bruner, Goodnow & Austin, 1956; Klausmeier, Ghatala & Frayer, 1974).

Hierarchy. One concept or class may include another, e.g. 'woman' and 'mother'. The division of nouns into count and mass nouns is an example from grammar (see Chomsky, 1965).

Overlap. A given entity may be assigned to two or more conceptual classes which are not hierarchically related, e.g. 'woman' and 'doctor'. To take the example of nouns again, Chomsky (1965, p. 79) argues that such subdivisions as 'proper' and 'human' overlap each other as do their complementary subdivisions, 'common' and 'non-human'.

'Fuzziness' of conceptual boundaries. The boundaries of conceptual categories are not usually sharply defined. Another way of expressing this is to say that the confidence with which objects may be assigned to a given category varies. For example 'cottage' is a category which shades into 'house' or 'hut'; it may be difficult to decide in which of the three categories a given building belongs (see Rommetveit, 1968). The same is true of grammatical classes. A given word may belong in more than one grammatical category and its 'nounness' or 'verbness' may vary. This point is noted by Kiss (1972).

Polythesis. Most human concepts are polythetic (see Sneath & Sokal, 1973) which means that no particular attribute or combination of attributes need necessarily be present (or absent) for an object to belong to a given category.

In fact the notion of a polythetic category is to some extent ambiguous. A strong sense of polythesis can be recognized which applies to the process of abstraction and means that even amongst the set of objects from which a 'schema rule' is abstracted 'no single instance will necessarily follow the rule in all respects and no single aspect of the schema rule will necessarily apply to all instances' (Evans, 1967, p. 87). However, there is a weaker sense applying to the process of recognizing new objects not in the original set. This is the familiar fact that people have a great capacity for correct identification in spite of omissions or additions (or simultaneous omission and addition, namely substitution) of the attributes of an object. A good example of this

Table 1. *The data for program CLST*02

Attributes

Objects	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
a	*	*	*	*							*							*		
b	*	*	*	*									*			*				
c	*	*	*	*								*								
d	*	*	*	*										*					*	
e	*	*	*	*													*			
f	*	*	*	*	*	*						*								
g	*	*	*	*	*	*										*		*		
h			*	*	*	*					*									
i			*	*	*	*							*							
j			*	*	*	*									*					*
k			*	*	*	*									*				*	
l			*	*	*	*	*	*	*	*										*
m							*	*	*	*				*						
n							*	*	*	*		*								
o							*	*	*	*					*					
p							*	*	*	*								*		

ability is in 'cloze' procedure tasks in which missing letters or words in running text are to be supplied (see Miller & Friedman, 1958). As will be shown, the clustering and recognition programs can model the weaker sense of polythesis but not the stronger.

Weighting of attributes. Many methods in numerical taxonomy are criticized by taxonomists for not recognizing the intuitive fact that some attributes are more significant in the development of classes than others. The attributes 'fur', 'four legs' and 'tail' are weak determiners of the concept 'cat' while 'retractile claws' and certain features of the teeth are very strong.

The following section describes the clustering and recognition programs. Their relevance to these six features of human conceptual systems is discussed at relevant points.

The model – CLST02 and RKLS02

The overall function of CLST02 is to sift out commonly co-occurring sets of attributes and to code less frequent sets of attributes in terms of their more frequent subsets whenever that is possible. The sets which are sifted out I have termed 'maximally associated' sets of attributes (MA-sets). A maximally associated set is defined as a set of attributes which occurs in a *greater* number of objects than any superset of that set (not an equal or lesser number). MA-sets are related to but not the same as maximally linked (ML) sets (Jardine & Sibson, 1971).

To illustrate the definition, consider the set of attributes (4, 11) in Table 1. These two attributes occur together in objects a and h and in no other objects. But the set of attributes (3, 4, 11) also occurs in these two objects. Since there is no other attribute which is present in both objects then set (3, 4, 11) is an MA-set but set (4, 11) is not.

A point to note about this clustering process is that the frequency of co-occurrence of attributes depends on the frequency of occurrence of objects. In Table 1 each object type occurs only once but there is nothing to prevent any or all of them occurring two or more times. Such multiple occurrences will affect the conceptual structure formed by the program. It is assumed that the same is true of human concept formation.

For every MA-set of attributes there is a corresponding set of objects in which the given attributes occur. It may be remarked in passing that, as with certain other clustering methods, an 'inverse' analysis may be performed in which the roles of objects and attributes are interchanged. CLST02 seems to be unique amongst methods giving overlapping clusters in that both objects and attributes are clustered in both direct and inverse analyses and the clusters produced are the same in both analyses. (Lambert & Williams, 1962, and Tharu & Williams, 1966, have achieved approximate solutions in the case of methods producing non-overlapping clusters.) To avoid confusion the general term 'element' will be used for a MA-set of attributes. A 'minimal element' is a single attribute and a 'composite element' is an element containing more than one attribute.

The particular algorithm described here is probably only one of several possible ways of realizing the same process and makes no pretence at modelling neural processes in detail.

1. The input data is a set of 'attribute lists', i.e. lists of attributes, one for each object. The first step is to set up a $n \times n$ 'frequency matrix' (or equivalent linkage structure to save space) where n is the total number of attributes.

2. The cells of the matrix are filled with counts of the number of objects in which each pair of attributes co-occur.

3. The largest count is selected (or an arbitrary choice in the case of ties) and the corresponding pair of attributes is joined to form a new composite 'attribute'. This is added to each of the appropriate attribute lists, i.e. those lists in which the two constituent attributes co-occur. These constituent attributes are not then deleted as might be expected. This is because they need to be left free to combine with other attributes later.

4. Each new element is assigned to a new node in a data structure with links connecting this node to the pre-established nodes for its two constituents. If, subsequently, it becomes a constituent of some other elements then further links connect its node to the nodes for these elements. This data structure, illustrated in Fig. 2, is formed on identical principles to that described in detail previously (Wolff, 1975) except that forwards and backwards links are not distinguished.

5. The frequency matrix is increased by one row and one column for the new composite attribute and the appropriate counts, taken from the attribute lists, are entered into the new cells with the following exceptions:

(*a*) The two cells corresponding to the pairing of the new element with each of its two immediate constituents.

If one or both of the immediate constituents are themselves composites (which is the rule in later stages of processing) then they each represent a set

Table 2. *Results from CLST02 and RKLS02*

Element or node number (E)*	The two immediate constituents of each element (X and Y)		The set of objects associated with each element or node (O)	The set of attributes (minimal elements) associated with each element or node (A)	The number of objects associated with each element ($F(O)$)	The number of attributes associated with each element ($F(A)$)	The salience (S) of each element (× 1000) $S = \frac{F(O)\times F(A)\dagger}{T(O)\times T(A)}$	The intersection (I) of A with H where H is the set of attributes (7,8,9,12)	The confidence ($C\times 100$) for object with attributes (7,8,9,12)	The intersection (I) of A with H where H is the set of attributes (7,8,9,10,15,1)	The confidence ($C\times 100$) for object with attributes (7,8,9,10,15,1)
1	2	3	4	5	6	7	8	9	10	11	12
12	—	—	c,f,n	12	3	1	9	12	100	∅	0
14	—	—	d,m	14	2	1	6	∅	0	∅	0
15	—	—	j,k,o	15	3	1	9	∅	0	15	100
18	—	—	a,g,p	18	3	1	9	∅	0	∅	0
21	3	4	a,b,c,d,e,f,g,h,i,j,k,l	3,4	12	2	75	∅	0	∅	0
(22)	6	21	f,g,h,i,j,k,l	3,4,6	7	3	—	—	—	—	—
23	5	22	f,g,h,i,j,k,l	3,4,5,6	7	4	88	∅	0	∅	0
(24)	2	21	a,b,c,d,e,f,g	2,3,4	7	3	—	—	—	—	—
25	1	24	a,b,c,d,e,f,g	1,2,3,4	7	4	88	∅	0	1	100
(26)	9	10	l,m,n,o,p	9,10	5	2	—	—	—	—	—
(27)	8	26	l,m,n,o,p	8,9,10	5	3	—	—	—	—	—
28	7	27	l,m,n,o,p	7,8,9,10	5	4	88	7,8,9	100	7,8,9,10	100
29	23	25	f,g	1,2,3,4,5,6	2	6	38	∅	0	1	29
30	20	23	j,l	3,4,5,6,20	2	5	31	∅	0	∅	0
31	19	21	d,k	3,4,19	2	3	19	∅		∅	0
32	18	25	a,g	1,2,3,4,18	2	5	31	∅		1	29
33	16	25	b,g	1,2,3,4,16	2	5	31	∅	0	1	29
34	15	23	j,k	3,4,5,6,15	2	5	31	∅	0	15	67
35	13	21	h,i	3,4,13	2	3	19	∅	0	∅	0
36	12	25	c,f	1,2,3,4,12	2	5	31	12	67	1	29
37	11	21	a,h	3,4,11	2	3	19	∅	0	∅	0
38	33	35	b	1,2,3,4,13,16	1	6	19	∅	0	1	14
39	32	37	a	1,2,3,4,11,18	1	6	19	∅	0	1	14
(40)	32	33	g	1,2,3,4,16,18	1	6	—	—	—	—	—
41	31	34	k	3,4,5,6,15,19	1	6	19	∅	0	15	33
42	30	34	j	3,4,5,6,15,20	1	6	19	∅	0	15	33
43	29	40	g	1,2,3,4,5,6,16,18	1	8	25	∅	0	1	14
44	29	36	f	1,2,3,4,5,6,12	1	7	22	12	33	1	14
45	28	30	l	3,4,5,6,7,8,9,10,20	1	9	28	7,8,9	20	7,8,9,10	20
(46)	25	31	d	1,2,3,4,19	1	5	—	—	—	—	—
47	23	37	h	3,4,5,6,11	1	5	16	∅	0	∅	0
48	23	35	i	3,4,5,6,13	1	5	16	∅	0	∅	0
49	18	28	p	7,8,9,10,18	1	5	16	7,8,9	20	7,8,9,10	20
50	17	25	e	1,2,3,4,17	1	5	16	∅	0	1	14
51	15	28	o	7,8,9,10,15	1	5	16	7,8,9	20	7,8,9,10,15	100
52	14	46	d	1,2,3,4,14,19	1	6	1	∅	0	1	14
53	14	28	m	7,8,9,10,14	1	5	16	7,8,9	20	7,8,9,10	20
54	12	28	n	7,8,9,10,12	1	5	16	7,8,9,12	100	7,8,9,10	20

* All minimal elements which are not MA-sets have been omitted.

† $F(O)$ and $F(A)$ are defined in columns 6 and 7. $T(O)$ and $T(A)$ are the total numbers of objects and attributes respectively (16 and 20 in Table 1).

of elements which includes the given main constituent, the constituents of that constituent, the constituents of those constituents, etc., down to the minimal elements. No counts are entered into the cells corresponding to pairings of the new element with all members of this set.

(*b*) In the same way, no counts are entered into the cells corresponding to all pairings of elements between the two sets represented by the two main constituents.

All the cells described in (*a*) and (*b*) are deleted or otherwise excluded from further consideration.

The procedure then returns to 3 and recycles repeatedly through 4, 5 and 3 until there are no cells left in the augmented frequency matrix containing counts greater than zero. The reason for pursuing the process down to the lowest frequencies is that it allows one to assign a structure to the attribute sets of single objects as will be shown. A point to note is that CLST02 can form associations between clusters which already have constituents in common. This is necessary to ensure that all MA-sets are found. An example appears in Fig. 2.

Illustrative result

The process may be illustrated with results from the artificial data shown in Table 1. These data are designed to cover points made in the introduction and to avoid the shortcoming in the 'frequency of co-occurrence' principle which has already been mentioned and which is discussed below.

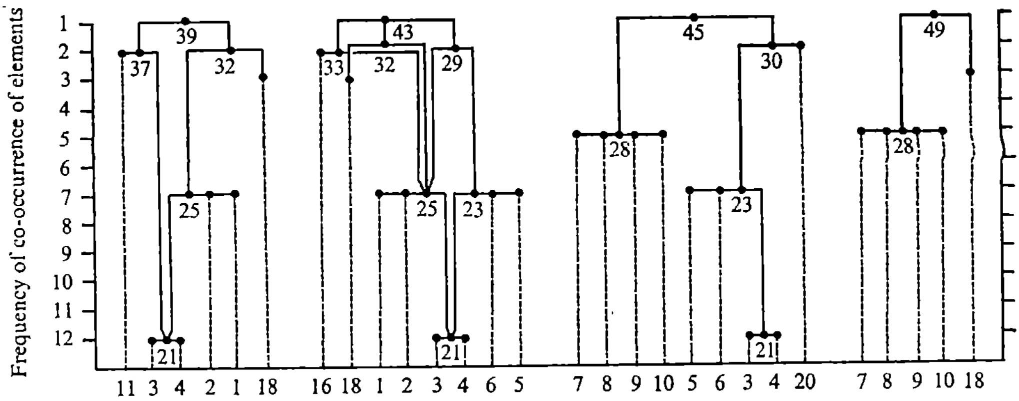
CLST02 reveals all the MA-sets in the data and these are shown in Table 2 (column 5) with the addition of six (marked with brackets in column 1) which are not MA-sets but are intermediate clusters formed by the program in the course of building up MA-sets. Only elements 12, 14, 15 and 18 of the 20 minimal elements turn out to be themselves MA-sets and only these four are shown. Elements 21–54, numbered in column 1 in order of formation, are all built out of the two constituents, X and Y, shown in columns 2 and 3. Using these two columns the complete structure of any element may be traced as will be seen in a moment. Column 4 (O) shows the objects corresponding to the attribute sets and columns 6 and 7 [$F(O)$ and $F(A)$] record the size of the object and attribute sets respectively.

It will be noticed that the later-formed MA-sets correspond to the attribute sets of the objects themselves. The interest in this apparently trivial result is in the structure assigned to these sets, a representative selection of which are shown in Fig. 1.

Each element is shown by a blob or 'node' in the structure, the height of which shows the frequency of co-occurrence of the elements dominated by that node (or, for minimal elements, simple frequency of occurrence). Strictly speaking all elements apart from the minimal elements have a binary structure as, for example, element 28 which is built up as (7(8(9, 10))). However, where more than two constituents join at the same frequency the element is effectively ternary, or, as in this case, quaternary. Another point to note is that the structures shown in Fig. 1 are actually interlocked in a complete data structure a partial representation of which is shown in Fig. 2.

We are now in a position to see how this clustering process mirrors three of the properties of concepts outlined earlier:

Hierarchy. The hierarchical relation of clusters can be seen in Fig. 2 where



(a) Element 39 (b) Element 43 (c) Element 45 (d) Element 49

Fig. 1. Four selected structures developed by program CLST02 from the data in Table 1.

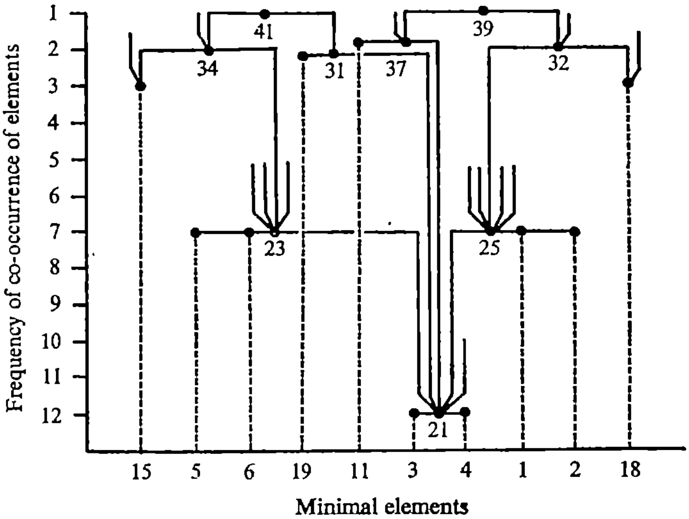


Fig. 2. Part of the data structure developed by program CLST02 from the data in Table 1.

element 21 is a shared constituent of elements 23 and 25. In terms of the corresponding object sets, element 21 is a superordinate category while elements 23 and 25 are subordinate to it.

Overlap. It may be seen from Table 2 that the sets of objects for elements 23 and 28 have object l in common even though these two elements are not hierarchically related. This then is a case of overlap. In a similar way objects f and g are common to the object sets for elements 23 and 35. This same overlap may be seen in Fig. 1(*b*) where nodes 23 and 25 are both dominated by node 29 which is itself dominated by node 43. These examples of hierarchy and overlap may also be seen in Table 1.

Salience. We can define a measure of salience for each cluster by:

$$S = \frac{F(O) \times F(A)}{T(O) \times T(A)},$$

where $F(O)$ and $F(A)$ have already been defined as the size of the object and attribute sets for the given cluster and $T(O)$ and $T(A)$ are the total numbers of objects and attributes respectively (i.e. 16 and 20 in Table 1). Values of S ($\times 1000$) are shown in column 8 of Table 2 and it can be seen that the four highest values (for elements 21, 23, 25 and 28) correspond to groupings of attributes which are intuitively the most prominent in Table 1. This very tentative measure of salience has the desirable property of varying from 0 to 1 but more work would need to be done to establish the extent to which it reflects the psychological salience of clusters.

THE RECOGNITION PROGRAM – RKLS02

Supplementary to the development of a classification system is identification or diagnosis, the process by which new objects are recognized or assigned to one or more of the pre-established categories. A computer program, RKLS02, has been written to model this process and is described here.

The data structure developed by CLST02 is, apart from the lack of distinction between backwards and forwards links, exactly the same as that developed by program MK10E (Wolff, 1975). To recapitulate, each element (including the minimal elements) is assigned to a node in the data structure. Every time two elements are joined to form a new composite element then a link is created between the new element and each of the immediate constituents – two links in all. The 'minor' links described for MK10E and retained in CLST02, are merely a programming device to allow several links to emerge from one node. In MK10E the recognition process has the strength of suggesting how context can influence recognition but it is implausible as a model of human recognition because it cannot handle any combination of contextual elements, it is a serial process and it cannot recognize a pattern containing distortions. These three defects are remedied in RKLS02.

This program takes as its input the data structure developed by CLST02 from data such as those in Table 1 and also a list of the attributes of a new object to be assigned to one or more of the concepts embodied in the data structure. This list of attributes need not correspond exactly to any of the original lists.

The first step is the assignment of each attribute to its corresponding minimal element or node in the data structure. Although this is a template matching process it is justified on the grounds that any recognition system, however sophisticated, must necessarily employ such a process at its most fine-grained level.

Next, signals are sent from each of these nodes up through the data structure to all nodes that can be reached from that node. At each one a trace or record is left of the attribute from which the signal originated. In this way lists of attributes are accumulated at some or all nodes. This process of sending signals through the network may loosely be termed parallel processing because several signals may be travelling in the network at the same time. It is perhaps more accurate to call it independence processing because the signals are independent of each other and although they *may* travel simultaneously they need not; indeed, when the process is modelled on a serial computer they cannot.

If the set of minimal elements or attributes dominated by any given node is designated as A and the set of attributes of the new object is H then the list of attributes accumulated at each node is simply the intersection, I, of A with H ($I = A \cap H$).

The set I for a given node may be found also at nodes having a higher $F(O)$. If, of all nodes with the same I, the one with the highest $F(O)$ is designated by the subscript j and the given node (which may be the same one) by the subscript i then a measure of 'confidence' that the new object belongs to the given node is taken to be simply:

$$C_i = \frac{F_i(O)}{F_j(O)}.$$

If set I is null then C_i is taken to be zero.

Strictly speaking, C is a conditional probability. It is the probability that the new object belongs in category i given that it contains the attribute set I. These 'probabilities' (shown in columns 10 and 12 of Table 2) do not add up to 100 per cent both because different categories have different sets I and also because some categories embrace or are embraced by others.

Imagine a new object having only attributes 7, 8, 9 and 12 (corresponding to the attributes of object n but with the omission of attribute 10). The sets I for this object are shown in column 9 of Table 2 and values of C in column 10. Most of the sets I are null and the corresponding nodes are not indicated in any way. Node 28, however, embraces three attributes and since no node having the same set I has a higher $F(O)$ then the object is assigned to category 28 with a C of 100 per cent. Since the set A for this category is attributes 7, 8, 9 and 10 the program has in effect made an inductive prediction on the basis of past evidence that, given attributes 7, 8 and 9, attribute 10 should also be present in the object. The justification for induction, especially when it produces wrong predictions as in this case, is the province of philosophers (e.g. Blackburn, 1973); the fact remains that inductive processes figure prominently in human cognition – the expectation that a set of attributes should be found together in the future given that they have always occurred together in the past is one that most people would acknowledge as plausible.

In a similar way category 54, which is subordinate to category 28, is indicated with 100 per cent confidence in spite of the fact that one of its attributes is missing. The other four categories which are subordinate to category 28 (45, 49, 51 and 53) are each indicated with 20 per cent confidence. Since the five categories are mutually exclusive only one can be chosen and this is necessarily category 54, the one with the highest confidence.

Rather than the new object being the same as one of the original objects with the omission of one or more attributes it may be the same as one of the original objects but with the addition of one or more attributes. The simplest possibility is that these supernumerary attributes are themselves new and therefore unrecognizable. In this situation they are simply filtered out at the first stage in the recognition process which then proceeds normally using only the attributes from the original set. More usually the 'noise' which distorts a pattern is itself composed of part and whole patterns. In the case of speech it may be other speech superimposed; in printed text it might be added or substituted letters. With print it is easy to see that there is a kind of conflict between those letters (attributes) which indicate the correct pattern and those which are noise. Usually, as for example in PSYCHBOLOGY, the true pattern is easily seen but in other cases such as CAUT it is not possible without additional clues to say whether one is dealing with CAT or CUT.

As an example, using our original data, imagine an object having attributes 7, 8, 9, 10, 15 and 1. The sets I and values of C are shown in columns 11 and 12 of Table 2. As one might expect the object is assigned to category 51 with a high degree of confidence, higher than for any other category having any or all of attributes 7, 8, 9, 10 or 15 in its set A. The effect of attribute 1 which is found in a relatively wide variety of categories is to determine most of these rather weakly with the single exception of category 25 which, like category 51, is indicated with a confidence of 100 per cent. This example is comparable with the CAT/CUT example given above. It is arguable that in PSYCHBOLOGY the B determines with 100 per cent probability the large class of words containing a B but this class is of such low salience, if it is recognized at all, that it seems to be completely discounted or ignored. However, category 25, although it covers a relatively wide range of subordinate categories, has been assigned a rather high salience and cannot be discounted in the same way. It seems that this aspect of the working of RKLS02 is rather less satisfactory than its handling of omissions.

Having described the recognition system we can now consider how the last three proposed properties of conceptual/recognition systems are modelled:

'Fuzziness' of concept boundaries. This means that, for a given category, it is possible to find objects which are assignable to that category with less than complete confidence. This is clearly true of the model and to that extent this property of conceptual systems is covered.

Polythesis. Perhaps the strongest feature of this model is to suggest how omissions of attributes from patterns may be filled in on an inductive basis. It meets the weak polythetic requirement in that any subset of the set of attributes referable to a category may be used to predict the remaining attributes. Supernumerary attributes are also handled but less successfully.

It remains true, however, that the clusters formed are in a sense 'defined' both intensionally as a set of attributes and extensionally as a set of objects and such definition is inconsistent with the strong sense of a polythetic category (see Jardine, 1969, who prefers the term 'family resemblance concept').

Weighting of attributes. Fig. 1(*d*) shows the structure of element 49 corresponding to object p. It should be clear from this example that the presence of attribute 18 in an object is more powerful in identifying element 49 (C = 33 per cent) than any of attributes 7, 8, 9 or 10 (C = 20 per cent). Such differences in weighting result naturally from the way the clustering process works without *ad hoc* provision.

GENERAL DISCUSSION AND CONCLUSION

We have seen that six significant properties of human conceptual/recognition systems are modelled with varying degrees of success. The other requirement, in accordance with schema theory, is that the clustering process should represent an efficient coding of the data. In order to assess the efficiency of coding we may assume that the coded data of Table 1 are stored as a set of nodes each node consisting of two or more 'pointers' pointing to the constituents of that node. Such a system would allow all nodes to be 'unpacked' to reproduce the sets of attributes of Table 1.

Table 3. *Hypothetical data discussed in the text*

		Attributes				
		1	2	3	4	5
Objects	1		*	*	*	*
	2	*		*	*	*
	3	*	*		*	*
	4	*	*	*		*
	5	*	*	*	*	

The minimum number of such pointers required in this case is 68 (2 × the number of composite elements) pointing to a maximum of 54 different nodes. So the minimum number of binary digits required for each pointer is $\log_2 54 \simeq 6$ and the total storage space required is $6 \times 68 = 408$ bits.

The data of Table 1 are in fact presented in only 320 two-state cells but this is only possible on the assumption that no attribute can occur more than once in any given object. If this assumption is abandoned then the simplest system for recording the data is a set of pointers, each signifying one attribute. For this data 94 such pointers would be required each needing $\log_2 20 \simeq 5$ bits of space. So the total storage space required is $5 \times 94 = 470$ bits. This is rather more than the figure of 408 bits computed above so it is clear that the clustering process does allow a modest economy in storage requirements in this case. If the number of objects was, say, 1000 but the number of basic patterns was still only 16 then dramatic economies become possible.

However if CLST02 is applied to the data shown in Table 3 no less than 30 MA-sets are isolated and the storage requirement calculated as before is about 350 bits. This compares most unfavourably with the simple pointer system requiring only about 60 bits. Such a set of objects could easily occur in our experience and it is most unlikely that we would employ such an inefficient coding system in this case. In fact it is intuitively clear that the most efficient coding would be achieved by assigning all five attributes to one node and coding the attributes of each object as the complete set *minus* one attribute.

Another way of expressing this problem is to say that the coding system employed in CLST02 is asymmetrical in the sense that it can only code an attribute set as one set with the *addition* of another rather than being able to code it either as one set with the addition of another or as one set with the *subtraction* of another. We may use the term 'positive coding' for the addition of one attribute set to another and 'negative coding' for the subtraction of one set from another and it seems that both kinds of coding need to be available to effect economical storage of all types of data. It should be clear from the above example of negative coding that such a system would from time to time form 'dummy' nodes or concepts having attribute sets never found complete in any one object. Likewise no single attribute from such a set need be found in all the objects belonging to the concept. Such concepts would be polythetic in the strong sense.

This work was conducted under the supervision of Dr Godfrey Harrison of University College, Cardiff. I am very grateful to him and also to Dr John Wilson of U.C.C. for useful discussions and for constructive comments on earlier drafts of this paper.

REFERENCES

ATTNEAVE, F. (1954). Some informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193.

ATTNEAVE, F. (1957). Transfer of experience with a class-schema to identification-learning of patterns and shapes. *J. exp. Psychol.* **54**, 81–88.

BERSTED, C. T. (1970). A general model of probabilistic concept formation. Unpublished doctoral dissertation, Texas Christian University.

BLACKBURN, S. (1973). *Reason and Prediction*. Cambridge: Cambridge University Press.

BOULTON, D. M. & WALLACE, C. S. (1970). A program for numerical classification. *Comput. J.* **13**, 63–69.

BROWN, B. R. & EVANS, S. H. (1970). Further application of the random adaptive module (RAM) system to schema theory. Technical Memorandum 4-70, Human Engineering Laboratories, US Army Aberdeen Research and Development Centre, Aberdeen Proving Ground, Md.

BRUNER, J. S., GOODNOW, J. J. & AUSTIN, G. A. (1956). *A Study of Thinking*. New York: Wiley.

CHOMSKY, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: M.I.T. Press.

COLE, A. J. (ed.) (1969). *Numerical Taxonomy*. London: Academic Press.

EVANS, S. H. (1964). A model for perceptual category formation. Part I. Unpublished doctoral dissertation, Texas Christian University.

EVANS, S. H. (1967). A brief statement of schema theory. *Psychon. Sci.* **8**, 87–88.

EVANS, S. H. & ELLIS, A. M. (1972). Annotated bibliography of reports on schema theory and related research. US Army Human Engineering Laboratories Technical Note, 1972 (Sept.).

JARDINE, N. (1969). A logical basis for biological classification. *Syst. Zool.* **18**, 37–52.

JARDINE, N. & SIBSON, R. (1971). *Mathematical Taxonomy*. New York: Wiley.

KISS, G. R. (1972). *Grammatical Word Classes. A Learning Process and its Simulation*. Edinburgh: MRC Speech and Communication Unit.

KLAUSMEIER, H. J., GHATALA, E. S. & FRAYER, D. A. (1974). *Conceptual Learning and Development. A Cognitive View*. London: Academic Press.

LAMBERT, J. M. & WILLIAMS, W. T. (1962). Multivariate methods in plant ecology. IV. Nodal analysis. *J. Ecol.* **50**, 775–802.

MCQUITTY, L. L. (1956). Agreement analysis: classifying persons by predominant patterns of responses. *Br. J. statist. Psychol.* **9**, 5–16.

MILLER, G. A. & FRIEDMAN, E. A. (1958). The reconstruction of mutilated English texts. *Inf. Control* **1**, 38–55.

NEEDHAM, R. M. (1963). A method for using computers in information classification. In C. M. Popplewell (ed.), *Information Processing 1962: Proceedings International Federation for Information Processing Congress, Munich*. Amsterdam: North-Holland.

OLDFIELD, R. C. (1954). Memory mechanisms and the theory of schemata. *Br. J. Psychol.* **45**, 14–23.

OLIVIER, D. C. (1968). Stochastic grammars and language acquisition mechanisms. Unpublished doctoral dissertation, Harvard University.

POSNER, M. I. & KEELE, S. W. (1968). On the genesis of abstract ideas. *J. exp. Psychol.* **77**, 353–363.

PREMACK, D. (1970). A functional analysis of language. *J. exp. Analysis Behav.* **14**, 107–125.

ROMMETVEIT, R. (1968). *Words, Meanings and Messages*. London: Academic Press.

ROSCH, E. H. (1973). Natural categories. *Cognit. Psychol.* **4**, 328–350.

SNEATH, P. H. A. & SOKAL, R. R. (1973). *Principles of Numerical Taxonomy*. San Francisco: Freeman.

SPÄRCK JONES, K. & JACKSON, D. M. (1967). Current approaches to classification and clump-finding at the Cambridge Language Research Unit. *Comput. J.* **10**, 29–37.

SPÄRCK JONES, K. & JACKSON, D. M. (1970). Use of automatically-obtained keyword classifications for information retrieval. *Inf. Storage* **5**, 175.

THARU, J. & WILLIAMS, W. T. (1966). Concentration of entries in binary arrays. *Nature, Lond.* **210**, 549.

UTTLEY, A. M. (1959). The design of conditional probability computers. *Inf. Control* **2**, 1–24.

UTTLEY, A. M. (1970). The informon: a network for adaptive pattern recognition. *J. theor. Biol.* **27**, 31–67.

WALLACE, C. S. & BOULTON, D. M. (1968). An information measure for classification. *Comput. J.* **11**, 185–194.

WOLFF, J. G. (1975). An algorithm for the segmentation of an artificial language analogue. *Br. J. Psychol.* **66**, 79–90.

ZAHN, C. T. (1971). Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE. Trans. comput.* **20**, 68–86.

(*Manuscript received* 11 *October* 1974; *revised manuscript received* 26 *May* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 391–397
Printed in Great Britain

QUALITATIVE ANALYSIS OF THE PERFORMANCE OF INTROVERTS AND EXTRAVERTS ON STANDARD PROGRESSIVE MATRICES*

By VIDHU MOHAN AND DALIP KUMAR
Department of Psychology, Panjab University, Chandigarh

The present study was designed to bring out the differences between extraverts and introverts on Standard Progressive Matrices. A group of 100 subjects of equal number of sexes were given a qualitative analysis of their performance on SPM. The study brought out the initial superiority of extraverts over introverts. But with passage of time and increase in difficulty level the introverts did significantly better as shown both by the t-ratios and the F-ratios. Personality interacted with sex to yield differences in performance of wrong scores only. These findings were discussed in terms of arousal, cortical excitation and reactive inhibition.

Does measurement of intelligence through a consolidated score imply that two or more subjects obtaining the same score are also undergoing the same mental process? According to Furneaux (1960), Guilford (1966) and Eysenck (1967) intellect can be classified three dimensionally thus: (*a*) 'mental speed', 'persistence', and 'error checking mechanism' (Furneaux, 1960); (*b*) 'operations', 'contents' and 'products' (Guilford, 1966); and (*c*) 'mental processes', 'test material' and 'quality' (Eysenck, 1967). These three-dimensional models of intellect, which were primarily based on deductions of first-order or second-order factor analysis, never implied that two subjects who obtained the same total score on an intelligence test were undergoing identical mental processes. Eysenck (1967), too, said that the assumption that 'equal scores are equivalent is unwarranted in the absence of proof and consideration of typical intelligence test papers'.

Eysenck (1967) suggested that these differences in the functioning of intellect might be related to his personality dimensions of Neuroticism (N), and Extraversion/Introversion (E/I). Discussing the overall relationship between intelligence and neuroticism, Eysenck (1967) quotes Farley's (1966) work where a non-linear relationship was found, persons average on N being superior in performance of intelligence tests to those with high or low scores on N. But the relation of E/I and intelligence is found to be linear and negative (Lynn & Gordon, 1961; Child, 1964; Madan, 1967; Eysenck, 1967). Introverts are supposed to opt for accuracy and extraverts for speed. But is it just the opting for speed or accuracy which is the determining feature of differences between extraverts and introverts on intelligence tests? Or is there something more in the understanding of the dynamics which can better be understood through a 'qualitative analysis' of the responses made on the test in terms of the items done rightly, wrongly, abandoned or not attempted, at various levels of task difficulty and time spent? Yates (1963) suggested an improvement in the items of SPM† by elimination of some and increase in time limit. But he did not ascribe the obtained differences in performance on SPM to any temperamental variable. The present study attempts to investigate such qualitative differences between extraverts and introverts on an intelligence test.

---

* Paper read at the Diamond Jubilee Session of Indian Science Congress, January 1973, Chandigarh, India.

† Standard Progressive Matrices will be written as SPM.

METHOD

A 2 × 2 design was adopted for the present study. There were two personality groups – introverts and extraverts, and two sexes – male and female. A total of 100 subjects were divided equally in four groups, making 25 subjects in each cell.

For the purpose of selecting subjects, the EPI (Eysenck & Eysenck, 1964) was administered to 500 Masters' students at Panjab University. Subjects scoring $M \pm \frac{1}{2}$ S.D. on E/I were categorized as extraverts and introverts. The mean scores on E/I of these extraverts and introverts were 14·3 and 6·5 respectively. The subjects ranged between 18 and 25 years in age.

The SPM test (Raven, 1960) was individually administered to the selected 100 subjects in five sets, A–E, in order of progressive difficulty. The total time taken by each subject on the test was also recorded.

RESULTS

A fourfold analysis was made of the performance of subjects on each of the five sets of SPM. The responses were scored in terms of the total number of right (R), wrong (W), abandoned (A), and non-attempted (NA) items in each of the five sets A, B, C, D and E. The means and S.D.s for each of these four response categories for (i) introverts, (ii) extraverts, (iii) males, (iv) females along with the respective *t*-ratios are presented in Table 1. The average time taken by each group and the statistical significance of the difference between means is presented alongside.

A graphic presentation of the differential performance of extraverts and introverts in each of the four response categories as a function of passage of time and increase in difficulty levels, as represented through the five sets of SPM, is given in Fig. 1.

The means and graphic presentation of performance on SPM of different personality groups, do reveal obvious differences. Hence a more extensive analysis is made to give statistical weightage to the observed differences. Since the performance of each subject is recorded on the five sets of SPM in four qualitatively different response categories, a repeated measure analysis of variance design is used (Edwards, 1968) of 2 × 2 × 5. There are two personality groups, two sexes and five repeated measures, i.e. the scores on five sets of SPM, A, B, C, D and E. Three such analyses of variance are calculated – one for the right responses, second for the wrong responses and in the third category the abandoned and not attempted responses were combined because of the similarity in trends of performance of the two personality groups as seen in Fig. 1. The results of the main effects and their interactions for right, wrong and combined scores of abandoned and not attempted items are presented in Table 2. The mean performance of extraverts and introverts and males and females for right and wrong scores in five sets is presented in Table 1. For the combined scores of abandoned and not attempted responses the respective means in five sets are 0·79, 1·06, 1·12, 0·74, 0·88 for introverts and, for extraverts, 0·12, 0·40, 1·40, 2·68 and 3·78. For males and females, 0·52, 0·78, 1·32, 1·76, 2·50 and 0·50, 0·68, 1·20, 1·66 and 2·16, respectively.

Table 1. *Means,* S.D.*s and t-ratios for introverts and extraverts and males and females on right, wrong, abandoned and not attempted scores of five sets of SPM and total time taken*

	Introverts		Extraverts			Males		Females		
	Means	S.D.s	Means	S.D.s	*t*-ratio	Means	S.D.s	Means	S.D.s	*t*-ratio
Set A										
R	10·44	1·15	11·22	0·95	3·70**	10·80	1·20	10·86	1·05	0·26
W	0·66	0·87	0·66	0·92	0·00	0·68	0·89	0·64	0·90	0·22
A	0·64	0·90	0·04	0·20	4·61**	0·36	0·69	0·32	0 74	0·28
NA	0·26	0·56	0·08	0·34	1·93	0·16	0·47	0·18	0·48	0·21
Set B										
R	10·38	1·19	10·64	1·29	1·05	10·60	1·37	10·42	1·11	0·72
W	0·56	0·79	0·96	0·99	2·24*	0·62	0·88	0·90	0·93	1·55
A	0·58	0·73	0·34	0·69	1·69	0·46	0·76	0·46	0·68	0·00
NA	0·48	0·74	0·06	0·24	3·84**	0·32	0·62	0·22	0·55	0·86
Set C										
R	9·36	1·48	9·00	1·60	1·16	8·90	1·62	9·46	1·43	1·83
W	1·52	1·66	1·60	1·28	0·27	1·78	1·76	1·34	1·08	1·50
A	0·68	0·84	0·92	0·97	1·32	0·80	0·93	0·80	0·90	0·00
NA	0·44	0·76	0·48	0·74	0·27	0·52	0·81	0·40	0·67	0·80
Set D										
R	9·02	1·49	7·26	1·77	5·38**	8·34	1·83	7·94	1·88	1·08
W	2·24	1·61	2·06	1·61	0·56	1·90	1·61	2·40	1·58	1·57
A	0·22	0·51	1·46	1·13	7·09**	0·90	1·11	0·78	1·04	0·56
NA	0·52	1·11	1·22	1·06	3·23**	0·86	1·23	0·88	1·15	0·09
Set E										
R	7·88	1·65	6·08	2·13	4·73**	6·88	2·11	7·08	2·11	0·47
W	3·22	1·82	2·14	1·54	3·20**	2·62	2·01	2·74	1·50	0·34
A	0·30	0·58	1·84	1·23	7·98**	1·04	1·24	1·10	1·23	0 24
NA	0·60	0·93	1·94	1·46	5·47**	1·46	1·49	1·08	1·27	1·37
Time taken	43·96	9·08	46·70	6·37	1·75	45·50	9·60	45·16	5·89	0·21

* Significant at 0·05 level.
** Significant at 0·01 level.

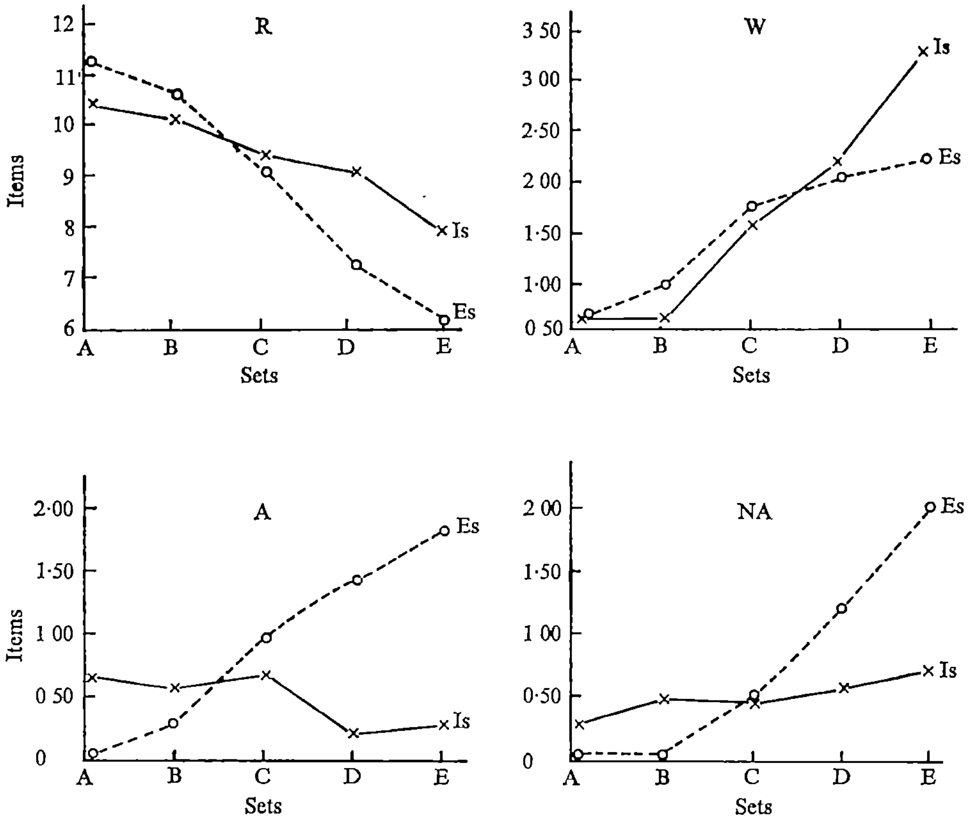


Fig. 1. Performance of introverts and extraverts on SPM.

Discussion

All the three analyses of variance (see Table 2) yielded highly significant F-ratios for the differences between the performance of subjects on the five sets of SPM. The means of the right, wrong (see Table 1) and combined abandoned–not attempted responses bear out that as one proceeds from set A to E the performance declines steadily, as would be expected from the progressive increase in difficulty level of the test. The results of the present work are perhaps best depicted in Fig. 1. The extraverts have an initial edge of superior performance on SPM over the introverts. But the latter seem to be steady and 'late developers' as far as performance on a continuous task is concerned. With passage of time and increase in the difficulty level, both the groups show a decline in the number of right responses and a corresponding increase in the wrong ones. But the ultimate performance decrement (as seen by the right responses) is marked and definite for the extraverts. This is also seen through the significant increase in the 'abandoned' (A) and 'not attempted' (NA) responses of the extraverts. The number of wrong items increases significantly more in the case of introverts as set E is reached (see Table 1). This overall trend is substantially borne out by the highly significant F-ratios of 23·18, 5·19 and 47·13 for the right scores, wrong scores and abandoned–not attempted scores respectively for the interaction

Table 2. *Analysis of variance for right, wrong and combined not attempted and abandoned scores for two personality, two sexes and five sets of SPM*

Source of variance	d.f.	Right scores: Sum of squares	Right scores: Mean squares	Right scores: *F*-ratio	Wrong scores: Sum of squares	Wrong scores: Mean squares	Wrong scores: *F*-ratio	Not attempted and abandoned scores: Sum of squares	Not attempted and abandoned scores: Mean squares	Not attempted and abandoned scores: *F*-ratio
Personality	1	41·472	41·472	7·905**	3·042	3·042	1·030	67·712	67·712	28·655**
Sex	1	0·288	0·228	0·005	0·882	0·882	0·029	2·312	2·312	0·982
Personality × sex	1	18·432	18·432	3·511	18·818	18·818	6·372*	0·000	0·000	0·000
Error (a) between subjects	96	503·700	5·246		283·536	2·953		225·544	2·349	
Sets	4	1039·448	259·987	175·920**	305·248	76·312	50·942**	217·928	54·482	38·804**
Personality × sets	4	137·108	34·277	23·175**	31·088	7·772	5·188**	264·688	66·172	47·131**
Sex × sets	4	13·452	3·363	2·273	12·568	3·142	2·096	1·448	0·362	0·257
Personality × sex × sets	4	7·348	1·837	1·242	12·392	3·098	2·068	3·760	0·940	0·668
Error (b) within subjects	384	568·060	1·479		575·504	1·498		539·176	1·404	
Total	499	2329·808			1243·078			1322·568		

* Significant at 0·05 level.
** Significant at 0·01 level.

Table 3. *Contingency table I*

Sets	Right scores E	Right scores I	Wrong scores E	Wrong scores I	Abandoned and not attempted scores E	Abandoned and not attempted scores I
A	11·22	10·44	0·66	0·66	0·12	0·90
B	10·64	10·38	0·96	0·56	0·40	1·06
C	9·00	9·36	1·60	1·52	1·40	1·12
D	7·26	9·02	2·06	2·24	2·68	0·74
E	6·08	7·88	2·14	3·24	3·78	0·88

between personality and sets. A contingency table of the means of the interaction in the three category response analysis is given in Table 3.

The significant F-ratio of 5·19 for personality × sets in the case of wrong scores and details of this interaction in Table 3 taken together with Fig. 1, show that the introverts made more errors than did the extraverts. As the introverts attempted significantly more items than did the extraverts over the same length of time (see results on abandoned–not attempted scores) they might thus be expected to make more errors. With the passage of time and increasing difficulty the extraverts seem to give up more often. This in a way contradicts Elliott's (1972) contention that 'high levels of task-difficulty may be expected to evoke "conditioned avoidance drive" (CAD) which has been learned after having previously been produced in failure situations (Jones, 1960)... Failure or non-receipt of expected reward is just as likely as punishment to evoke such a response, which would adversely affect introverts' learning efficiency.' Or perhaps CAD might not have been formed in the present situation of SPM because failure was not simultaneously perceived. Eysenck (1967), on the other hand, attributes the better performance of the introverts to their high arousal and the poorer performance of extraverts to their poor arousal. In this connexion he quotes the results of better 'short recall times' for extraverts and 'long recall times' for introverts in paired associate learning. This higher arousal leading to better consolidation (Eysenck, 1967) plus the quicker cortical excitation slower accumulation of I_r and its faster dissipation (Eysenck, 1952) account more for the present results where, in a progressively difficult and long task, the introverts become superior in performance towards the last sets of SPM.

In a continuous task like the SPM, the extraverts have an advantageous start, but start dropping in their performance with the passage of time. Madan (1967) and Lynn & Gordon (1961) too had explained the better performance of introverts on SPM and in academic attainment as due to greater persistence and desire for accuracy. In the present experiment the introverts required less time to complete the test than did the extraverts; however, the difference between the means fails to reach any level of significance (see Table 1). The results on accuracy are more conclusive and indicate that in this group the introverts performed better than did the extraverts.

The present findings show no differences whatsoever on the basis of sex, in the two groups. However, the variable of sex interacts with personality to yield significant F-ratio of 6·37 (see Table 2) in the case of wrong scores only. Table 4 presents a contingency table to show the respective mean interactions in the case of wrong scores. Table 4 shows that male extraverts committed the least errors closely followed

Table 4. *Contingency Table II*

	Wrong scores	
	Male	Female
Extraverts	1·24	1·72
Introverts	1·79	1·48

by female introverts and the maximum number of errors were committed by male introverts.

In the light of the present results a suggestion could be made concerning the comparison between extraverts and introverts on intelligence tests or performance. Firstly, the length of the test material should be shortened sufficiently to reach the optimum for both groups. Secondly, if this shortening is not feasible, a disinhibitor or an optimum rest pause should be interposed to arrest the rapid decline in the performance of extraverts due to accumulation of I_r. Mohan (1968), Mohan & Mohan (1976) found the introduction of a disinhibitor at an opportune time in a continuous task like the pursuit rotor to affect the extraverts advantageously. And finally the sex differences should be kept in mind.

REFERENCES

CHILD, D. (1964). The relationships between introversion extraversion, neuroticism and performance in school examinations. *Br. J. educ. Psychol.* **34**, 187–196.

EDWARDS, A. L. (1968). *Experimental Design in Psychological Research.* New York: Holt Rinehart & Winston.

ELLIOTT, C. D. (1972). Personality factors and scholastic attainment. *Br. J. educ. Psychol.* **42**, 23–32.

EYSENCK, H. J. (1952). *Scientific Study of Personality.* New York: Macmillan.

EYSENCK, H. J. (1967). Intelligence assessment: A theoretical and experimental approach. *Br. J. educ. Psychol.* **37**, 81–98.

EYSENCK, H. J. & EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory.* London: London University Press.

FARLEY, F. H. (1966). Individual differences in solution time in error-free problem solving. *Br. J. soc. clin. Psychol.* **5**, 306–309.

FURNEAUX, W. D. (1960). Intellectual abilities and problem solving behaviour. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology.* London: Pitman.

GUILFORD, J. P. (1966). Intelligence: 1965 model. *Am. Psychol.* **21**, 20–26.

JONES, H. G. (1960). Learning and abnormal behaviour. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology.* London: Pitman.

LYNN, R. & GORDON, I. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Br. J. educ. Psychol.* **31**, 194–203.

MADAN, V. (1967). The relation of neuroticism and extraversion to intelligence and educational attainment at different age levels. PhD thesis, Panjab University, Chandigarh.

MOHAN, J. (1968). Disinhibition and reminiscence. *Br. J. Psychol.* **59** (3), 277–279.

MOHAN, J. & MOHAN, V. (1976). Personality determinants of the effect of disinhibitor on reminiscence. *Psychologica Belgica* (in the Press).

RAVEN, J. C. (1960). *Guide to the Standard Progressive Matrices; Sets A, B, C, D and E.* (Revised.) London: Lewis.

YATES, A. J. (1963). A further study of Progressive Matrices (1947). *Br. J. educ. Psychol.* **33**, 307–312.

(*Manuscript received* 12 *September* 1974; *revised manuscript received* 20 *June* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 399–406
Printed in Great Britain

COGNITIVE STRATEGIES AND MULTIDIMENSIONAL SCALING

BY C. O. FRASER*
University of Canterbury, Christchurch, New Zealand

The problems involved in equating multidimensional scaling solutions with cognitive strategies are considered. Empirical evidence is presented to suggest that scaling solutions are relatively insensitive to the combination rule used by subjects, and thus may be used as an indication of the basic variables, but not their method of combination, that underlie a cognitive strategy.

The increasing use of multidimensional scaling (MDS) as a descriptive and explanatory aid to studying the behaviour of individuals to complex stimuli raises some important theoretical questions regarding the interpretation that can be placed on such solutions. Two clearly divergent viewpoints have emerged regarding this. One is to view MDS as a quantitative psychological theory, an attempt to model the basic psychological processes underlying a set of cognitive judgements. The alternative, however, is to regard MDS merely as a way of organizing and summarizing such information, a sort of descriptive statistic for a large mass of data.

The theoretical arguments for both sides of this controversy have been adequately detailed elsewhere and will not be repeated here (see, for example, Beals, Krantz & Tversky, 1968; Arabie & Boorman, 1973). However, neither of these take full account of the practicalities of actually performing an MDS, the effect of the limitations of available algorithms on theoretical considerations. A major problem is that the weak restrictions imposed by non-metric MDS often require combining observations over more than one individual in order to extract a solution, a procedure with obvious difficulties for a psychological model interpretation. One can minimize this to some extent, for example by using algorithms allowing individual transformations between data and distances (e.g. POLYCON II, Young, 1973), or different weights or saliences on a common set of dimensions (e.g. INDSCAL, Carroll & Chang, 1970). However, we shall need empirical evidence to show to what extent such a solution does correspond to the strategies used by subjects to generate their dissimilarities.

Tversky & Krantz (1969) showed that for one class of stimuli at least (schematic faces), the individual dissimilarities corresponded closely with the ordinal constraints predicted by a dimensional representation model, i.e. one in which independent dimensional differences contribute additively to the overall dissimilarity.

However, it does not follow that this will hold for all possible classes of stimuli. Rapoport & Fillenbaum (1972), for example, collected proximity data for verbs in the 'have' family, and showed that a hierarchical clustering provided a much better indication of subjective strategies than a multidimensional interpretation. Miller (1969) has argued that in general a multidimensional interpretation is inappropriate for stimuli seen by the subject to have a hierarchical organization.

The present experiment concerns another class of stimuli which has caused difficulties for the dimensional model, namely circular or directional stimuli. Arabie &

---

* Requests for reprints should be sent to C. O. Fraser, who is now at Gippsland Institute of Advanced Education, PO Box 42, Churchill, Victoria, 3842, Australia.

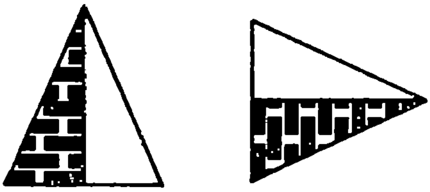


Fig. 1. A typical stimulus pair. The stimuli are related by the transformations g_1 through 90° and g_2.

Boorman, for example, cite the difficulty of interpreting the dimensions of the colour circle separately as an argument against the insistence on a dimensional representation.

An unpublished study by Gregson (1972) investigated some aspects of the similarities of directional stimuli and suggested that they may offer particular problems for MDS. Gregson used isosceles triangles with one half white and the other black (see Fig. 1). These varied in either their angle of rotation, i.e. the orientation of the centre line, or in the side of the triangle which was coloured black, which can also be regarded as a rotation about the centre line. Gregson regarded this stimulus series as merely representing a set of closed group transformations, corresponding to the rotations used to generate the series. Following Gregson's notation we will use g_1^k to denote rotation of the centre line in the plane of the triangle through k°. Similarly rotation of the triangle about the centre line through 180° is denoted by g_2^{180} or, simply g_2 as the incremental angle is not in this case arbitrary. The periodicity of the group is defined as the number of transformations required to convert a stimulus back to itself. Thus g_1^k has periodicity $360°/k$, and g_2 has periodicity 2.

Any stimulus can thus be specified by some combination of g_1 and g_2 from an arbitrary starting point or, alternatively, they can be used to specify the transformations required to convert one stimulus into another. Gregson hypothesized that the cognitive processes used by subjects to evaluate similarities of circular stimuli could be represented by a mapping on to one or more such group transformations. However, as he points out, their subjective strategies need not be confined to these transformations in a simple manner. For example, he suggests a third transformation (g_3), a rotation of the triangle 180° about the base line (i.e. a flip-over operation) which may be used in preference to the equivalent combination of g_1 through 180° and g_2.

Gregson regressed similarity against angular difference to test for the periodicity of best fit. He found that some subjects had a better fit for a periodicity of less than 360° (i.e. maximum dissimilarity at 180°). He attributed this to the use of the g_3 rotation defined above, giving a false impression of a periodicity 180° due to the higher similarity at exactly 180°.

However, an alternative more direct explanation is possible for the better fit at exactly 180°. Subjects may, instead of measuring the angular difference between the directed centre lines, measure only the difference between the undirected lines, producing a transformation of periodicity 180°. Both of these strategies could be used to produce a mixed similarity function of θ, the angular difference, and 2θ, the undirected difference. A simple regression of the type used by Gregson would not distinguish between this mixed strategy and the combination of g_1 and g_3 which he suggests.

Unless the subject has grossly misunderstood the instructions, his similarity function must be some function of the two basic transformations g_1 and g_2 which generated the series. Thus when we desire to extract the basis on which the similarity function is predicated we are not interested in extracting only g_1 and g_2, as these are common to all, but the idiosyncratic way in which they are combined. For example, it may provide a simpler and more realistic indication of the subject's strategy to describe the function in terms of some other set of transformations (such as Gregson's g_1 and g_3) even though they are reducible to the basic transformations.

EXPERIMENT I

A conventional paired comparisons similarity estimation experiment was run on 17 subjects using the stimulus series from the Gregson (1972) experiment. Twelve stimuli comprising a representative sampling from the rotation and Black/White reversal series were used, giving 66 pairs in all. Each pair was projected onto a 1 m × 1 m screen for a constant duration of about 8 sec, and was rated on a scale of 0–100. Examples of different and identical stimulus pairs were presented during the instructions; subjects were told to score identical stimuli as 100, but were not told any number to pair with any non-identical pair. All subjects received all combinations of the 12 stimuli used. The subjects were first year undergraduate psychology students.

Instructions to subjects

The following instructions were read to subjects prior to the experiment: 'You will be shown pairs of patterns side by side on the screen, and you are to give a number between 0 and 100 to each pair of patterns. This number is to represent a point on a scale expressing how similar the two patterns are as you see them. "100" means they are the same, "0" means they are not at all similar. Any number between 0 and 100 is possible. You call out your answers as soon as you can.

'I will show you some examples of pairs of patterns and the numbers you should give them: – These are the same and therefore should be scored 100, – these are not the same and therefore should be scored some number less than 100.

'Apart from that it is over to you to decide how to allocate the numbers . . .'

The subjects then received some other instructions regarding reaction time and confidence measurements which will not be considered here.

RESULTS

The 17 individual-subject similarity matrices were first analysed by INDSCAL, a fully metric MDS program which places heavy emphasis on a dimensional representation and allows for individual differences in the relative importance of the dimensions (Carroll & Chang, 1970).

The solution was interpretable in up to five dimensions; after this they began to repeat themselves. Some of the two-dimensional plots of the five-dimensional solution are shown in Figs. 2–5 below, with the dimensions numbered as extracted by INDSCAL. The first two dimensions clearly show the primary rotation dimension, the stimuli being arranged in their correct order around the circle (Fig. 2). Dimension 4 separates the stimuli into two groups which are completely independent of the rotation dimensions, as is shown by the two straight lines in Fig. 3. These groups correspond exactly to the equivalence classes defined by rotation about the centre line (labelled g_2^0 and g_2^1 for same and different respectively). The three-dimensional arrangement of dimensions 1, 2 and 4 would thus be a pair of circles in parallel planes, one above the other.

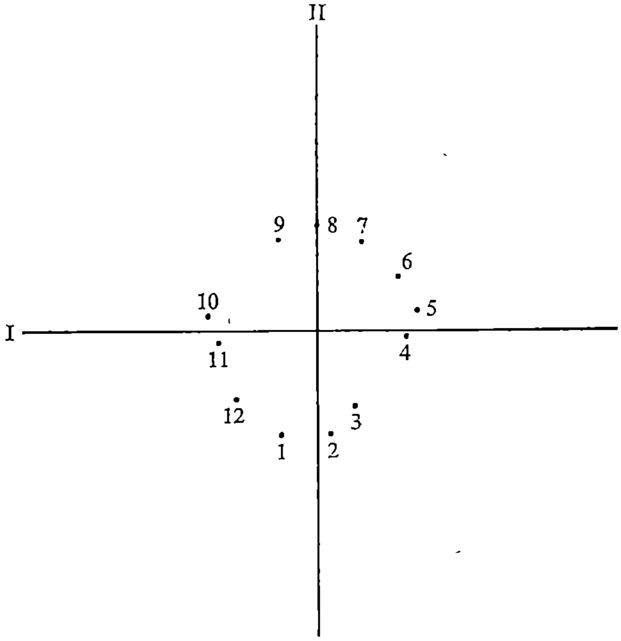


Fig. 2

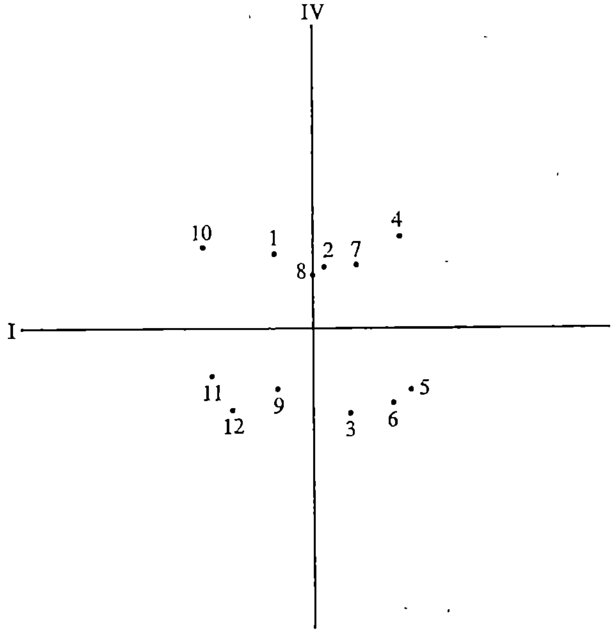


Fig. 3

For legend see facing page.

The next two dimensions (dimensions 3 and 5 in Indscal oddly enough) are obviously related to some sort of secondary rotation effect (Fig. 4). The irregularity of this figure however clearly offers an insufficient basis to distinguish a secondary rotation from the Gregson 'g_3' hypothesis which would also group opposing stimuli but with

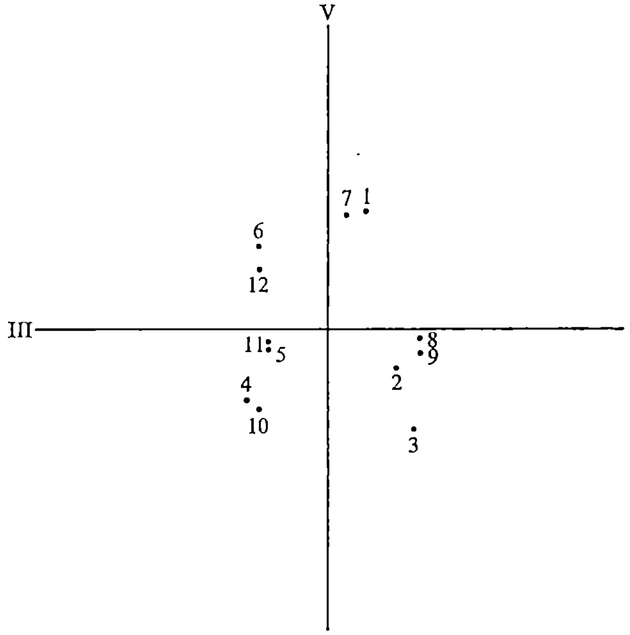


Fig. 4

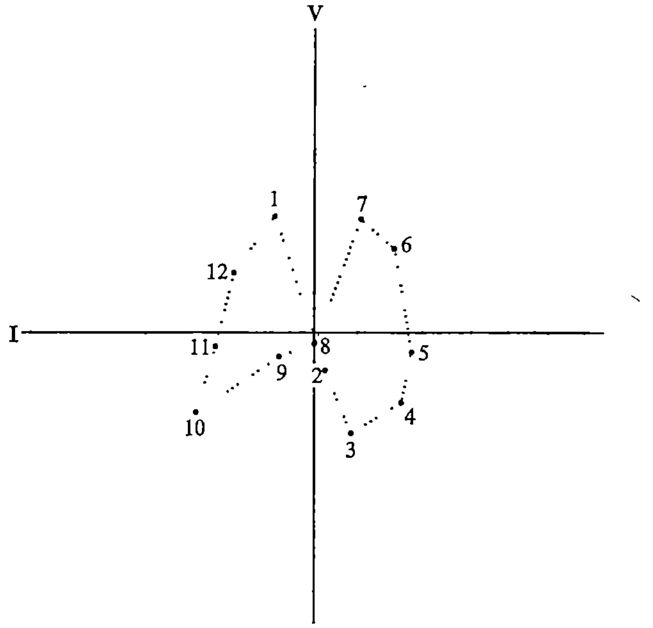


Fig. 5

Figs. 2–5. INDSCAL five-dimensional solutions. The stimuli are numbered in order of increasing angle of rotation.

the pairs not necessarily in order. The clue to this problem comes from examining some of the overlap plots, with one of the primary rotation dimensions against one of the secondary rotation dimensions. Fig. 5, the plot of dimension 1 vs. dimension 5, shows the stimuli in a distinctive figure-of-eight pattern. Students of electronics will recognize this as one of a class known as Lissajoux figures, which are created by using different sine waves to control the x and y coordinates of a cathode-ray display. As the relative frequency and phase of the two sine waves change so does the pattern, and thus uniquely determines the identity of the signal. For example, the circle in Fig. 2 is itself a Lissajoux figure for two sine waves of equal frequency with phase difference $\pm 90°$. Thus it is a graph of $\sin \theta$ vs. $\sin (\theta - 90°)$, i.e. $x = \sin \theta$ and $y = \cos \theta$, the parametric equations of the circle.

The figure-of-eight pattern obtained identifies two sine waves with frequency ratio 1:2 and a phase difference of zero. Thus it is a graph of $\sin \theta$ vs. $\sin 2\theta$, suggesting that dimensions 3 and 5 may be interpreted as $\cos 2\theta$ and $\sin 2\theta$, i.e. a secondary rotation with periodicity 180°.

The other three combinations of primary and secondary dimensions do indeed support this interpretation.

For example, dimension 2 vs. dimension 5 also showed a figure-of-eight, but in the opposite direction with a different crossover point. This corresponds to a phase difference of 180°, i.e. $\sin \theta$ vs. $\sin (2\theta + 180°)$ or $\sin \theta$ vs. $\sin 2(\theta + 90°)$, which, as the waves are continuous, is equivalent to $\sin (\theta - 90°)$ vs. $\sin 2\theta$ or $\cos \theta$ vs. $\sin 2\theta$.

Individual strategies

Having now obtained a group scaling solution that appears intuitively plausible, it now remains to be shown that a set of weightings on these five dimensions do in fact give a reliable indication of individual subjects' strategies. Since we have only two dimensions, one of them two-valued, we can in this case directly identify these strategies by examining the similarity function, i.e. by plotting the mean similarity for each angular difference, separately for g_2^0 and g_2^1.

A series of linear regressions of similarity over angular difference with periodicities 360°, 180° and 90° were also carried out to provide a more quantitative comparison. A few cases of the strategies found are described below by way of example.

Subject 1

The g_2^0 graph showed a U-shaped function with a minimum at about 144°, indicating a primary rotation effect (periodicity 360°) with a fairly strong secondary rotation effect (periodicity 180°). However, the g_2^1 graph had a sawtooth pattern, indicating either a periodicity of 90° or some sort of 90° effect analagous to g_3, i.e. a single transformation, say g_4, which corresponds to rotation about the line 45° to the centre lines of both triangles.

Subject 3

This graph was U-shaped for g_2^1 and monotonically decreasing for g_2^0, indicating periodicities of 180° and 360° respectively.

Subject 4

Both graphs showed perfect linear decreasing relationships, indicating a primary rotation effect and little else.

Many other subjects also showed major divergences from the INDSCAL derived strategies, with a typical result being a difference in periodicity between g_2^0 and g_1^2

Table 1. *Individual strategies:* INDSCAL *weights vs. regression significance levels*

	INDSCAL			Regression analysis					
				g_2^0			g_2^1		
Subject	g_1 (360°)	g_1 (180°)	g_2	g_1 (360°)	g_1 (180°)	g_1 (90°)	g_1 (360°)	g_1 (180°)	g_1 (90°)
1	0·24	0·34	0·33	—	*	—	—	—	—
2	0·12	0·00	0·89	—	—	—	—	—	—
3	0·46	0·31	0·03	*	—	*	**	*	—
4	0·62	0·00	0·10	***	—	—	***	—	—
5	0·62	0·00	0·10	***	—	—	***	—	—
6	0·63	0·02	0·03	***	—	—	***	—	—
7	0·37	0·47	0·15	—	**	—	**	*	—
8	0·47	0·20	0·06	*	—	—	**	—	—
9	0·28	0·41	0·08	—	**	—	—	—	—
10	0·26	0·17	0·70	**	—	—	—	—	**
11	0·51	0·00	0·16	**	*	**	**	—	—
12	0·48	0·29	0·06	—	**	—	***	—	—
13	0·15	0·54	0·07	—	**	—	—	—	**
14	0·60	−0·07	0·19	***	—	—	***	—	—
15	0·37	0·27	0·12	**	—	—	—	—	—
16	0·34	0·42	0·16	—	**	—	**	—	—
17	0·51	0·07	0·22	***	*	—	**	—	*

* $P < 0{\cdot}05$; ** $P < 0{\cdot}01$; *** $P < < 0{\cdot}01$.

stimulus pairs. The findings from both analyses are summarized in Table 1 above. This shows the significance levels of the regression analyses, and the weightings given the three strategies identified by INDSCAL (averaged over two dimensions where appropriate). Only in the cases of the four subjects with pure strategies (S4, S5, S6, S14) is there a reasonable matching on both sides of the table.

DISCUSSION

The similarity functions used by most subjects would be clearly inadmissible under the Beals, Krantz & Tversky conditions for dimensional representation, as the differences along one dimension are a function of the differences along another. Nevertheless, the INDSCAL solution has successfully retrieved the major transformations known to underlie the data, in spite of using a metric that does not even approximate the way they are combined.

This severely weakens the argument insisting on treating MDS as a psychological model of behaviour, as a successful data reduction has been achieved in spite of using an inappropriate model. This apparent robustness of the scaling solution to the combination rule used by subjects of course means that the converse is also true; if we are satisfied that the dimensions of a configuration correspond to the variables used by subjects, this does not guarantee that they are combined in the same manner as the distance function used to derive the configuration.

REFERENCES

ARABIE, P. & BOORMAN, S. A. (1973). Multidimensional scaling of measures of distances on partitions. *J. math. Psychol.* **10**, 148–203.

BEALS, R., KRANTZ, D. H. & TVERSKY, A. (1968). Foundations of multidimensional scaling. *Psychol. Rev.* **75**, 127–142.

CARROLL, J. D. & CHANG, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalisation of Eckart-Young Decomposition. *Psychometrika*, **35**, 283–319.

GREGSON, R. A. M. (1972). Similarities within a set of closed group transformations. Unpublished report, University of Canterbury.

MILLER, G. A. (1969). Psychological method to investigate verbal concepts, *J. math. Psychol.* **6**, 169–191.

RAPOPORT, A. & FILLENBAUM, S. (1972). An experimental study of semantic structures. In Shepard, Romney & Nerlove (eds), *Multidimensional Scaling; Theory and Applications in the Behavioural Sciences*, vol. II. Seminar Press.

TVERSKY, A. & KRANTZ, D. H. (1969). Similarity of schematic faces – a test of interdimensional additivity. *Percept. Psychophys.* **5**, 124–128.

YOUNG, F. W. (1973). Conjoint scaling. Thurstone Psychometric Laboratory, Rep. No. 118.

(*Manuscript received* 12 *September* 1974; *revised manuscript received* 3 *July* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 407–412
Printed in Great Britain

RIGHT- AND LEFT-HANDED COUNTING OF BRAILLE DOTS IN SUBJECTS UNACCUSTOMED TO BRAILLE

BY D. H. MYERS
Salop Area Health Authority

Previous work has shown that children acquainted with Braille can usually read better with their left than with their right hand. It is also known that, at least in right-handed people, the right cerebral hemisphere is dominant for counting dots.

The present paper compares, in right-handed subjects unacquainted with Braille, the comparative skill of right and left middle (M) and index (I) fingers in counting Braille dots. In one experiment with 30 subjects no significant difference was found between LI and RI fingers. Another experiment failed to reveal any R/L differences although it detected a superiority of I over M fingers at $P < 0{\cdot}025$. When the task was split between the two I fingers working simultaneously they performed significantly ($P < 0{\cdot}001$) better than when working alone. This latter finding is thought to have a bearing on the failure to detect R/L differences; other possible explanations of the findings are also discussed.

The principle of cerebral dominance was first established for speech over a century ago. Since then it has been found to apply to several cerebral functions (Butler, 1971; Milner, 1971). In a recent series of investigations Kimura has found that in right-handed people the right hemisphere is better than the left at counting dots when these are presented tachistoscopically. Dots thus presented to the left peripheral visual field in normal subjects are more accurately counted than when presented to the right (Kimura, 1966). Furthermore the ability to count dots presented tachistoscopically is impaired more in epileptics who have had a right, than in those who have had a left temporal lobectomy (Kimura, 1963).

As well as providing an important insight into cerebral organization, these findings might have a practical bearing on the question of which hand is best for Braille reading. Does the right hemisphere dominance for counting dots confer any advantage on the left hand? Or are there other factors which might nullify this or even give the advantage to the right hand? One such factor is left cerebral dominance for language. As familiarity with Braille symbols grows and language becomes inextricably a part of their perception, so the left hemisphere probably becomes increasingly influential. There remains the possibility that the cerebral commissures transmit information between the hemispheres so effectively that cerebral dominance in practice confers no advantage on either hand. Hermelin & O'Connor (1971), however, have found that in a group of blind children the left hand is significantly superior to the right in reading Braille. The present study examines the issue in subjects who have had no previous experience of Braille.

METHOD

Subjects and apparatus

All subjects were right-handed young adults without scars or deformities on the fingers tested, and without any history of disease or trauma to the upper limbs. A subject was regarded as being right-handed if he or she used the right hand to throw, to write and to perform tasks requiring dexterity. Different sets of subjects were used for each experiment. Before examination the hands were dipped in warm water for three minutes, washed and carefully dried. For Expts. I and II, 48 separate groups of Braille dots were punched on a strip of Braille cardboard. The sequence of the groups along the strip was random. (The situation for Expt. III was slightly different and is described under that experiment.) The groups were presented by a rotating drum, one by one and at constant speed through an aperture to the finger under examination. After each presentation of a group of dots the subjects called out the number of dots they thought the group contained.

Experimental design

The middle and index fingers of the right and left hands were examined in Expts. II and III. Latin squares were used to isolate and measure variance between subjects and variance between the order in which each finger was tested. It was anticipated that the effect of testing any one finger might, because of practice, influence the score of the next finger tested. The Latin squares were therefore balanced for such carry over effects by ensuring that the testing of every finger was preceded equally often by the testing of each of the other fingers. Three 4 × 4 Latin squares were used for each experiment. The randomization was carried out in such a way that order effects were isolated for each experiment as a whole and not separately for each Latin square. The design was taken from Cochran & Cox (1957).

Expt. I was conducted as above, except in two details. Instead of being confined to comparing single fingers it compared the index fingers used singly with their use simultaneously. Since 30 subjects were examined ten 3 × 3 Latin squares were used.

Several of the scores – the percentage of the 48 groups correctly counted – were over 80 per cent. For this reason the scores were converted to angular measure by the arcsin transformation before analysis of variance. The mean scores in the 'results' section are quoted in percentages: the figures in the tables are in angular measure.

RESULTS

Experiment I

This had two objectives. The first was to compare the effectiveness of left and right index fingers in counting groups of dots when the position of the dots within each group was random. The second was to determine if counting was easier when two fingers were used rather than one. For the first objective a tape was prepared (tape A) with each group of dots containing between one and five dots. Each dot within a group was randomly allocated to one of the nine positions provided by a 3 × 3 dot Braille printing grid. For the second objective two tapes were prepared (the B tapes) such that if the dots of one were superimposed on those of the other, a pattern of dots identical to that of tape A was obtained. This was done by splitting each group of dots into two subgroups by a straight line whose position was determined by random numbers. One of the subgroups thus obtained was printed on one tape B and the other subgroup on the other tape B. These three tapes were mounted on three drums on the same axle.

Thirty subjects were examined. Each subject was tested in three ways: right index on tape A; left index on tape A; right index on one tape B and left index simultaneously on the other tape B with the subject adding the findings of the two fingers. Each group of dots was presented stationary for 2 sec during which the finger made exploratory movements. The mean scores obtained were: left index 55·10; right index 50·85; both used together 63·80. It is apparent from the analysis of variance (Table 1) that the use of both fingers together was significantly more effective than the use of either alone. It will also be noticed that there was a trend towards better left-hand performance but this failed to reach statistical significance.

Table 1. *Experiment I*

Source	SS	d.f.	MS	*F*	*P*
Subjects	2745·0	29	94·7	—	—
Order	33·8	2	16·9	—	—
Carry over	55·9	2	28·0	—	—
Between fingers	702·5	2	351·2	13·5	< 0·001
(I) Used singly vs. used together	630·2	1	630·2	24·3	< 0·001
(II) L index vs. R index	72·3	1	72·3	2·8	n.s.
Error	1400·4	54	25·9	—	—

Table 2. *Experiment II*

Source	SS	d.f.	MS	*F*	*P*
Subjects	2082·3	11	—	—	—
Order	50·0	3	—	—	—
Carry over	59·8	3	—	—	—
Between fingers	156·2	3	52·1	3·0	< 0·05
(I) Left vs. right	42·9	1	42·9	2·5	n.s.
(II) Middle vs. index	100·9	1	100·9	5·8	< 0·025
Error	470·5	27	17·4	—	—

Table 3. *Experiment III*

Source	SS	d.f.	MS	*P*
Subjects	1857·7	11	—	—
Order	61·3	3	—	—
Carry over	135·2	3	—	—
Between fingers	36·0	3	12·0	n.s.
(I) Left vs. right	9·5	1	9·5	n.s.
(II) Middle vs. index	22·8	1	22·8	n.s.
Error	583·7	27	21·6	

Experiment II

The two objectives of this experiment were: firstly, to determine if there was any difference in the skill of the right and left hands when counting patterns of dots (as opposed to the randomly disposed dots of Expt. I); and, secondly, to compare the skill of the middle with the index fingers on this task. Twelve subjects were examined.

The 48 groups on the Braille cardboard strip now consisted of a random sequence of the Braille symbols for A, K, L, X and Y (which symbols consist of one to five dots respectively). Each symbol was presented stationary for 1·3 sec, during which time the finger made exploratory movements. The mean scores obtained were: right index 68·10, right middle 61·20; left index 69·55; left middle 66·35. It will be seen from Table 2 that the index fingers were superior to the middle fingers but there is no significant difference between the fingers of the right hand and those of the left.

Experiment III

This was identical to Expt. II except in two circumstances. Firstly, only the letters A to X were used. Secondly, the fingers did not make exploratory movements but remained still while the drum drove the symbols under them at a speed of 5 cm/sec there being a pause of 1·5 sec between each delivery of a symbol. Under these conditions the mean scores were: right index 70·30; right middle 68·90; left index 69·70; left middle 68·05. Table 3 shows that there is no significant difference between right fingers compared with left fingers or between the index fingers compared with the middle fingers.

DISCUSSION

Several investigators have examined the possibility that the tactile sensitivity of the two sides of the body might differ. Weinstein & Sersen (1961) tested the palms, forearms and soles, and Fennell, Satz & Wise (1967) the palms, of normal subjects. Carmon (1971) in the course of investigating patients with unilateral cerebral lesions examined the palmar tactile sensitivity of a group of control patients without neurological disorder. None of these investigators was able to find statistically significant R/L differences in right- or left-handed groups of normal (i.e. without neurological disability) subjects.

The first point to make about the present study is the obvious one: the experiments do not show that there is no difference in the skill of the left and right hands in counting Braille dots; they have merely failed to reveal that there is such a difference. This may be because the difference is very small and would require many more subjects to demonstrate it. This point can be made in a comparative sense by observing that if there is a genuine R/L difference then Expt. II suggests that it is very probably a good deal smaller than the difference between the index and middle fingers. Alternatively, the failure to reveal a difference may even have occurred in the face of a fairly substantial difference if for purely chance reasons the sample of subjects used did not typify the R/L balance in the population. Finally, however, the findings may result from the fact that there really is no difference between the two hands. This conclusion would have to be reconciled with Hermelin & O'Connor's finding of left-hand superiority. In the first place, the tasks in this investigation and in theirs differed: their subjects read Braille and this clearly involves skills in addition to counting dots, skills for which there may be right cerebral dominance. Secondly, differences in the skill of right and left hands in those who have learned Braille may be acquired rather than innate. If the innate skill of the two hands is equally balanced, it is clear that any extraneous factor which favours one of the hands will, by reinforcement with practice, soon make that hand the more skilful. One such extraneous factor considered by Hermelin & O'Connor is the convenience of starting with the left hand because Braille reading is done from left to right. To get a rough idea of the extent to which this convenience might favour the use of the left hand the following was done:

Thirty-seven right-handed subjects were asked – using a pretext – to use any finger they liked to feel Braille symbols – firstly at the left-, and then at the right-hand margin of a 10 in strip of Braille material placed before them. This material was placed in the same position as a 10 in wide book the subject had earlier placed in front of him as for reading. Twenty-two subjects used the left hand for the left margin; this appeared to be for convenience because they used the right hand for the right margin.

There are several possibilities to account for the failure to find any difference in the skill of the hands in counting dots despite right cerebral dominance for this process. One is that any purely sensory advantage the left hand might enjoy is nullified by the more effective exploratory movements of the right fingers in right-handed subjects. However, Expt. III, by failing to find left-hand superiority when this factor was removed, makes this an unconvincing explanation. Another possibility is raised by the work of Kaufman, Lord, Reese & Volkmann (1949). They found evidence of at least two different methods of counting dots which they call 'subitizing' and 'estimating'. It cannot be excluded that these two processes are dominant in different sides of the brain. Their relative contribution to counting by sight (the modality in which Kimura demonstrated right-sided dominance) and to counting by touch may differ, and to an extent which obscures the right dominance effect when touch is used for testing. Another, and more plausible, explanation for the findings is that touch is too crude a sensation to put the perception of number under sufficient load to reveal any subtle degree of cerebral asymmetry. In support of this explanation is the finding that an index finger used alone is not able to put to anything like full stretch the ability of the cortex to establish a count from tactile information. This is apparent from the superiority shown in Expt. I of both index fingers used together compared with either alone.

This absence of evidence of cerebral dominance in at least one aspect of the hands' sensory ability is in marked contrast to the obvious manifestation of dominance in their movement. This may result from some important differences between the organization of the most highly developed forms of movement and of sensation. The hands, in which the brain's capacity to organize movement finds its most delicate and versatile expression, can operate independently of one another, and each is only effectively represented in one half of the brain. These circumstances would enable any slight initial advantage due to cerebral dominance to become rapidly reinforced and entrenched by practice. But the most highly developed sense organs – the ears and the hemimaculae of the eyes – cannot function independently of their fellows and are bilaterally represented in the brain. There is no way, therefore, in which any initial asymmetry due to cerebral dominance can be fostered by the preferential use of one ear or of one pair of conjugate hemimaculae. This latter point is relevant to the ability of the fingers to count dots because, except in those blind from very early years, the cerebral organization for the perception of dots was presumably established largely, if not wholly, in conjunction with the development of vision.

The experiments reported here explore only a very limited aspect of Braille reading and inferences must be tentative. With this proviso, Expt. II suggests that the habit of preferring the index fingers to the middle when learning Braille is not an arbitrary matter but based on their innate superiority. It does not seem to matter which index finger is used, however, and there are practical reasons for educating both: subsequent damage to one would then be less of a calamity; and as the first experiment suggests, their simultaneous use might at least in some instances increase capacity.

I should like to thank those who acted as subjects. I should also like to thank Mrs M. Jones, Miss Gibson-Smith, Mr A. Laurie and the Staff at the RNCB Rowton Castle for their advice and interest. I am most grateful to Electronics and Applied Physics Division, AERE, Harwell, for the generous loan of the machine used in the investigation. The West Midlands Regional Health Authority kindly defrayed incidental costs. Dr O'Connor generously commented on the manuscript.

Br. J. Psychol. (1976), 67, 3, *pp.* 413–417
Printed in Great Britain

THE BEHAVIOURAL EFFECTS OF SINGLE AND COMBINED STRESSORS: A TEST OF AROUSAL THEORY

BY P. M. LIND
Australian National University, Canberra

Arousal theory assumes that a single physiological dimension underlies the curvilinear (inverted-*U*) relation of behavioural efficiency to level of stimulation. Traditional tests of this assumption, which involved correlating different physiological measures of arousal, have produced equivocal results. The present study predicted that, if arousal is unitary, then stressors which separately induce it should be additive in their effects on behavioural efficiency. Although curvilinear changes in behavioural efficiency were found for white noise and shock separately, their combination failed to support the prediction.

Duffy (1967), Hebb (1955), Lindsley (1951) and Malmo (1959) have argued that a single intensity dimension, so-called 'arousal', underlies changes in behavioural efficiency. Specifically, arousal is defined as a single intensity dimension that is 'indicated roughly by any one of a number of physiological measures (e.g. skin resistance, muscle tension, EEG, cardiovascular measures and others)' (Duffy, 1962, p. 5). This dimension is assumed to underlie the curvilinear (inverted-*U*) changes in behavioural efficiency which have often been observed with increases in stimulation (e.g. Yerkes & Dodson, 1908; Hockey, 1970).

The assumption that arousal is unitary implies there will be substantial correlations among different physiological measures. In general, there have been two types of correlation studies. In the first, inter-individual studies have investigated whether, for example, an individual's heart rate can be predicted from a knowledge that skin conductance is high relative to some group of people. Although such studies have typically found low physiological correlations, both Duffy (1967) and Malmo (1959) have argued that this type of correlation is neither a sensitive nor appropriate measure of arousal. In the second type of approach, intra-individual studies have investigated whether, for example, an individual's heart rate and skin conductance will systematically increase with increases in stimulation. These studies have produced equivocal results. Whereas Malmo & Bélanger (1967) have reviewed numerous experiments which reported high correlations, Lacey (1967) reviewed other experiments which found low ones.

Because this line of research has not resolved the question of whether arousal is unitary, the aim of the present study is to investigate this problem from a somewhat different orientation. Specifically, it is hypothesized that, if arousal is unidimensional, then stimuli which induce arousal will be additive in their effects on behavioural efficiency. For example, if increases in two stimuli administered separately produce curvilinear changes in behavioural efficiency, then increases in these two stimuli presented together should produce monotonic changes in behavioural efficiency. At low levels of stimulation, the two together should produce more arousal and improve performance; at moderate levels of stimulation, the two together will exceed optimal arousal and impair performance; and, at high levels of stimulation, the two together will produce more arousal and impair performance even further. These predictions were tested by examining changes in behavioural efficiency as a function of white noise intensity, shock intensity and the simultaneous presentation of white noise and shock.

METHOD

Subjects

Subjects were 200 naïve male albino rats obtained from the John Curtin School of Medical Research, Canberra, Australia. They were between 100 and 120 days old at the beginning of the experiments and were maintained on *ad lib.* food and water.

Apparatus

The apparatus was a shuttlebox, consisting of two 12 × 7 × 10 in. compartments covered throughout with stainless steel and separated by a motor-driven, stainless-steel door. The raising and lowering of the door took approximately 0·1 sec. The ceiling of each compartment was covered with 1 in. Plexiglass. In the centre of each compartment at ceiling level, two 1·5 W clear light bulbs served as the CS. The floor of each compartment consisted of $\frac{1}{8}$ in. stainless-steel rods spaced 1 in. apart. Scrambled electric shock was delivered by a Grass 24KR Constant Current Stimulator. In the centre of each compartment 1 in. above the floor level, a photocell provided the recording of response latencies. At the side of the shuttlebox, a 5 in. speaker was mounted and overlapped equally into the two compartments. The speaker was powered by a Scott 811-B Random Noise Generator and provided white noise of 1200–4800 c/sec. All stimulus presentations and timings were programmed by Digital K-Series Logic. Response latencies in 0·1 sec units were printed out by a Sodeco PL103 Counter.

Procedure

Following the habituation trials described by Theios, Lynch & Loew (1966), subjects received 100 test trials with a CS–UCS interval of 5 sec and an inter-trial interval of 20 sec. A trial consisted of opening the door and presenting the CS in the occupied compartment. If subject responded by running into the other compartment within 5 sec, the CS terminated. These responses were defined as avoidance responses. If subject failed to respond, the UCS came on until either he did leave the occupied compartment or 30 sec had elapsed. Such responses were defined as errors and terminated the CS and UCS.

Subjects were randomly assigned to groups and run in random order. Five groups, ten subjects in each, received white noise, either 75, 85, 95, 105 or 115 dB, as UCS; five groups, ten subjects in each, received shock, either 0·1, 0·4, 0·7, 1·0 or 1·3 mA, as UCS; and five groups, 20 subjects in each, received the simultaneous presentation of white noise and shock, either 75 dB + 0·1 mA, 85 dB + 0·4 mA, 95 dB + 0·7 mA, 105 dB + 1·0 mA or 115 dB + 1·3 mA, as UCS.

RESULTS

For the white noise groups, means and standard deviations of the percentage of trials on which subject avoided the UCS and the latency of these responses are given in Table 1. For per cent avoidance, an analysis of variance indicated that the quadratic trend (e.g. Winer, 1962) across groups was significant ($F = 6{\cdot}0$, d.f. = 1, 45, $P < 0{\cdot}05$). There were no other significant trends, nor were the between-groups differences significant ($F = 2{\cdot}0$, d.f. = 4, 45, $P > 0{\cdot}1$). For latency, where the raw scores violated the assumption of homogeneity of variances, the logarithmic transformation (e.g. Winer, 1962) was found to have no significant quadratic trend ($F = 3{\cdot}6$, d.f. = 1, 45, $P = 0{\cdot}07$), no other significant trends and no significant between-groups differences ($F = 1{\cdot}7$, d.f. = 4, 45, $P > 0{\cdot}1$).

For the shock groups, means and standard deviations of per cent avoidance and latency are given in Table 2. For per cent avoidance, there were a significant quadratic trend ($F = 6{\cdot}1$, d.f. = 1, 45, $P < 0{\cdot}05$), a significant linear trend ($F = 6{\cdot}3$, d.f. =

Table 1. *Means and standard deviations for per cent avoidance and for the latency of avoidance responses as a function of white noise intensity*

Measure (dB) ...	75	85	95	105	115
Per cent avoidance					
Means	33·1	50·6	46·5	51·4	38·2
Standard deviations	23·8	18·5	11·3	17·2	17·2
Avoidance latency					
Means	3·3	2·8	2·9	2·7	3·0
Standard deviations	0·9	0·5	0·3	0·5	0·3

Table 2. *Means and standard deviations for per cent avoidance and the latency of avoidance responses as a function of shock intensity*

Measure (mA) ...	0·1	0·4	0·7	1·0	1·3
Per cent avoidance					
Means	34·5	53·4	46·9	26·6	20·6
Standard deviations	13·4	18·9	28·2	23·1	22·4
Avoidance latency					
Means	3·3	2·8	2·3	2·4	2·9
Standard deviations	0·1	0·4	0·6	0·5	0·4

1, 45, $P < 0{\cdot}05$) and significant between-groups differences ($F = 4{\cdot}0$, d.f. = 4, 45, $P < 0{\cdot}05$). For latency, where the variances were not homogeneous, the Mann–Whitney U test indicated that the 0·7 mA and 1·0 mA groups did not differ ($U = 93{\cdot}5$) but the 0·7 mA group was significantly faster than the 0·1 mA ($U = 67$, $P < 0{\cdot}01$), 0·4 mA ($U = 75$, $P < 0{\cdot}05$) and 1·3 mA groups ($U = 74$, $P < 0{\cdot}05$) and, similarly, the 1·0 mA group was significantly faster than the 0·1 mA ($U = 65{\cdot}5$, $P < 0{\cdot}01$), 0·4 mA ($U = 74{\cdot}5$, $P < 0{\cdot}05$) and 1·3 mA groups ($U = 76$, $P < 0{\cdot}05$).

For the white noise plus shock groups, means and standard deviations of per cent avoidance and latency are given in Table 3. For per cent avoidance, there were no significant linear trend (d.f. = 1, 95, $F < 1$), a significant quadratic trend ($F = 5{\cdot}5$, d.f. = 1, 95, $P < 0{\cdot}05$), and no significant between-groups differences ($F = 1{\cdot}6$, d.f. = 4, 95, $P > 0{\cdot}1$). For latency, there were no significant linear trend (d.f. = 1, 95, $F < 1$), a significant quadratic trend ($F = 7{\cdot}6$, d.f. = 1, 95, $P < 0{\cdot}01$) and significant between-groups differences ($F = 2{\cdot}5$, d.f. = 4, 95, $P < 0{\cdot}05$).

An additional analysis was made to compare the white noise and shock groups with the white noise plus shock groups. Scores from the 75 dB group were added to scores from the 0·1 mA group and the mean of this combined group was compared with the 75 dB + 0·1 mA group and so on. For per cent avoidance, the results of planned, orthogonal contrasts for these five comparisons in order were $F = 1{\cdot}4$, d.f. = 1, 190, $P > 0{\cdot}1$; $F = 2{\cdot}5$, d.f. = 1, 190, $P > 0{\cdot}1$; d.f. = 1, 90, $F < 1$; d.f. = 1, 190, $F < 1$; d.f. = 1, 190, $F < 1$, respectively. For latency, where the variances were not homogeneous, the results of the Mann–Whitney U test in order were $U = 192{\cdot}5$, $P > 0{\cdot}1$; $U = 171{\cdot}5$, $P > 0{\cdot}1$; $U = 182{\cdot}5$, $P > 0{\cdot}1$; $U = 155$, $P > 0{\cdot}1$; and $U = 78{\cdot}5$, $P > 0{\cdot}1$, respectively.

Since per cent avoidance was lower at every intensity for combined stimuli than for the mean of the separate ones, a final analysis compared the white noise and shock functions with the white noise plus shock function. Based on Snedecor &

Table 3. *Means and standard deviations for per cent avoidance and the latency of avoidance responses as a function of the simultaneous presentation of white noise and shock*

Measure ...	75 dB + 0·1 mA	85 dB + 0·4 mA	95 dB + 0·7 mA	105 dB + 1·0 mA	115 dB + 1·3 mA
Per cent avoidance					
Means	25·3	40·3	40·6	32·9	27·0
Standard deviations	19·5	24·2	26·2	28·9	26·9
Avoidance latency					
Means	3·5	3·0	2·7	2·8	3·1
Standard deviations	0·7	0·7	0·8	0·8	1·1

Cochran's (1967) comparisons of linear combinations of means, a test was devised to examine the hypothesis that an addition of the linear coefficients of the separate functions did not differ from the linear coefficient of the combined function and, simultaneously, an addition of the quadratic coefficients of the separate functions did not differ from the quadratic coefficient of the combined function. That is, the hypothesis was $\beta_1+\beta_2-\beta_3 = 0$ and $\gamma_1+\gamma_2-\gamma_3 = 0$ where the linear coefficients of the white noise function, shock function and white noise plus shock function were $\beta_1 = 1{\cdot}1$, $\beta_2 = 5{\cdot}9$ and $\beta_3 = -0{\cdot}4$ respectively, and the quadratic coefficients of these functions were $\gamma_1 = -3{\cdot}7, \gamma_2 = -4{\cdot}5$ and $\gamma_3 = -3{\cdot}5$ respectively. This hypothesis was rejected at the 0·05 level (d.f. = 2, 118, $F = 3{\cdot}2$).

DISCUSSION

Two comparisons suggested that the behavioural effects of combined presentations of two stimuli do not differ from the behavioural effects of their separate presentation. Statistical trends, which described the relation of either the percentage or latency of avoidance responses to level of stimulation, were predominantly quadratic for both combined and separate stimuli. Furthermore, individual data points at all levels of combined intensities did not differ from the mean of individual data points at all levels of the separate ones. It should be noted that inter-subject variances seem reasonably comparable to those reported in other shuttle avoidance studies (e.g. Moyer & Korn, 1964). However, the results cannot be interpreted simply since per cent avoidance is lower at every intensity for combined stimuli than for the mean of the separate ones. In fact, the combined function did differ in its linear and quadratic coefficients from the separate functions. It is not clear, however, how this finding should be theoretically interpreted. Because current approaches to interactions between stressors do not predict this effect, it would seem wise to determine whether it will occur with other dependent variables. Until then, any firm conclusion that separate stressors do not add along the arousal dimension would appear incautious.

Also, two alternative interpretations of the results warrant some consideration. It might be suggested that physical systems do not respond in a linear fashion throughout the range of an impressed stimulus (e.g. Patton, 1970). In other words, 'ceiling' effects may have been appearing when white noise and shock were combined. This suggestion assumes that the two stimuli are additive over part of the range of stimulation and yet the present data found no such effect, even at low levels where the behavioural function had not reached its optimal or ceiling value. An alternative explanation which might be suggested is that subjects selectively attend to one stimulus and ignore the other. While this hypothesis seems plausible at low levels of stimulation, a mechanism of selective attention which switches off channels of potentially damaging input seems somewhat unadaptive. Furthermore, some attention theorists have admitted that animals can simultaneously process two stimuli (e.g. Mackintosh & Turner, 1971). Thus, it seems difficult to specify a mechanism by which either white noise or shock could be ignored without simply describing the results that two stimuli have the same effect as one.

In conclusion, the present study found a curvilinear relation of behavioural efficiency to level of white noise or shock separately; however, a predicted increase in arousal by combining them failed to alter this relation. Although no physiological measures were taken and so arousal theory was not directly tested (Duffy, 1967), it seems difficult to interpret the results in terms which warrant the assumption that arousal underlies behavioural changes.

This paper is based on a report to the Ninth Conference of the Australian Psychological Society, Perth, 1974. I am deeply grateful to Professor K. A. Provins, Dr W. P. Bellingham and L. H. Storlien for their invaluable supervision and to Dr David Chant for statistical advice.

REFERENCES

Duffy, E. (1962). *Activation and Behaviour.* New York: Wiley.

Duffy, E. (1967). The nature and development of the concept of activation. In R. Haber (ed.), *Current Research in Motivation*, pp. 278–281. New York: Holt, Rinehart & Winston.

Hebb, D. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.* **62**, 243–254.

Hockey, G. (1970). Effect of loud noise on attentional selectivity. *Q. Jl exp. Psychol.* **22**, 28–36.

Lacey, J. (1967). Somatic response patterning and stress: some revisions of activation theory. In M. H. Appley & R. Trumbull (eds), *Psychological Stress*, pp. 14–42. New York: Appleton-Century-Crofts.

Lindsley, D. (1951). Emotion. In S. Stevens (ed.), *Handbook of Experimental Psychology*, pp. 473–516. New York: Wiley.

Mackintosh, N. & Turner, C. (1971). Blocking as a function of novelty of CS and predictability of UCS. *Q. Jl exp. Psychol.* **23**, 359–366.

Malmo, R. (1959). Activation: a neuropsychological dimension. *Psychol. Rev.* **66**, 367–386.

Malmo, R. & Bélanger, D. (1967). Related physiological and behavioural changes: what are their determinants? *Research Publications, Association for Research in Nervous and Mental Disease*, **45**, 288–318.

Moyer, K. & Korn, J. (1964). Effect of UCS intensity on the acquisition and extinction of an avoidance response. *J. exp. Psychol.* **67**, 352–359.

Patton, G. (1970). Combined autonomic effects of concurrently applied stressors. *Psychophysiology* **6**, 707–715.

Snedecor, G. & Cochran, W. (1967). *Statistical Methods*, pp. 308–310. Iowa: Iowa State University Press.

Theios, J., Lynch, A. & Loew, W. (1966). Differential effects of shock intensity on one-way and shuttle avoidance conditioning. *J. exp. Psychol.* **72**, 294–299.

Winer, E. (1962). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

Yerkes, R. & Dodson, J. (1908). The relation of strength of stimulus to rapidity of habit-formation. *J. comp. neurol. Psychol.* **18**, 459–482.

(*Manuscript received* 24 *January* 1975; *revised manuscript received* 12 *July* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 419–427
Printed in Great Britain

THE INFLUENCE OF REVERSIBILITY AND TRUNCATION ON THE INTERPRETATION OF THE PASSIVE VOICE BY YOUNG CHILDREN

BY MARGARET HARRIS
Department of Psychology, Birkbeck College, University of London

Reversible and non-reversible active and passive sentences, half of which were truncated and half not, were presented to children aged between three and seven years. Comprehension of each sentence was tested by the simultaneous presentation of picture pairs, one of which showed the actor and object in the same roles and the other in the opposite roles to that described by the sentence. Selection of the pictures which the children considered to match the sentences confirmed the prediction that children whose comprehension of passive sentences was imperfect were aided not only by non-reversibility but also by agent-deletion. The most difficult type of passive to comprehend was one which was reversible and contained an agent. Non-reversibility was also found to aid comprehension of active sentences by the younger subjects.

Bever (1970) suggests that the passive is not correctly interpreted by four-year-old children because they adopt an incorrect strategy of assigning the invariable interpretation actor–action–object to any noun–verb–noun (NVN) sequence in a sentence. The generalization of this strategy, which correctly interprets actives and thus the majority of English sentences, leads to incorrect identification of the actor and object of passive sentences. However, semantic factors appear to influence the use of the NVN strategy from about the third year of life and McNeill (1970) summarizing Bever's (1970) results suggests that young children learn to adopt the following strategy: (1) assume that noun–verb–noun stands for actor–action–object (the NVN strategy); (2) if this is implausible, assume that noun–verb–noun stands for object–action–actor.

One factor that may affect the plausibility of interpretation is non-reversibility. Several studies (e.g. Turner & Rommetveit, 1967, 1968; Bever, 1970; Sinclair, Sinclair & De Marcellus, 1971; Strohner & Nelson, 1974) have supported Slobin's (1966) original finding that young children are able to utilize the semantic information provided by non-reversibility in order to interpret correctly passive sentences which on the basis of syntax alone they would be likely to misinterpret. For instance, application of the NVM strategy to sentence (1),

(1) The flowers are being watered by the girl

which appeared in Slobin's (1966) experiment, leads to the interpretation *The flowers are watering the girl* which is anomalous because it breaks the selection restrictions on the kind of subject and object that can occur with the verb 'to water'.

Herriot (1969) investigated another kind of non-reversibility using sentences which were non-reversible only in the sense that they described a probable relationship between two people; although their converse was also semantically possible. Sentence (2) is probable whereas sentence (3) is not.

(2) The lawyer advised the client.
(3) The client advised the lawyer.

The present experiment investigates non-reversibility of this latter kind.

Another factor which may influence the way in which young children comprehend the passive voice is truncation. Svartvik's (1966) extensive study of the passive voice in English showed that the majority of passive sentences do not contain an agent. On grounds of frequency alone, then, it is perhaps not surprising to find that the passive first appears in child speech in its truncated form, typically with the verb 'to get' rather than the more formal 'to be' as in *Blacky didn't get hit* (Menyuk, 1971). Moreover, if misinterpretation of the passive does arise from its confusion with the active form and the resulting overgeneralization of the NVN strategy, then agent-deleted passives should be correctly interpreted at an earlier age since they are less easily confused with actives. It was predicted that agent-deletion would serve as a syntactic cue to the inappropriateness of the NVN strategy just as non-reversibility provides a semantic cue.

Previous experiments have used either an acting-out technique in which the subject acts out active and passive sentences with two puppets (e.g. Bever, 1970; Cromer, 1970; Maratsos, 1974; Strohner & Nelson, 1974) or a choice between pairs of pictures the incorrect version of which depicts the actor and objects the wrong way round (e.g. Fraser, Bellugi & Brown, 1963; Gaer, 1969; Scholnick & Adams, 1973). With either technique the subject has a 50 per cent chance of choosing the correct interpretation. Moreover, these studies used a very small number of instances to test comprehension. Fraser *et al.* (1963), for example, used only two pairs of utterances to investigate several grammatical contrasts including the passive; and a recent replication of Bever's (1970) results was based on performance on three instances with a score of two correct being accepted as evidence for passive comprehension (Maratsos, 1974). In order to provide a less equivocal indication of the child's ability to understand different types of passive sentences, the Fraser *et al.* (1963) technique was extended so that 12 instances of each sentence type were presented to each child. The ability of a child to comprehend a particular sentence type could thus be determined by inspection of performance on 12 instances to see if it was significantly different from chance. If the number of successes was significantly above chance then the performance on that sentence type could be deemed a pass. All other levels of performance could be considered failures, varying from performance at chance level down to an above chance level of errors, which would suggest consistent use of an inappropriate strategy.

The purpose of the present experiment was to study the influence of both a syntactic cue (i.e. agent-deletion) and a semantic cue (i.e. non-reversibility) on the evolving comprehension of passive sentences by young children. It was predicted that three distinct patterns of performance would emerge from the three age groups to be studied. The youngest children were expected to find great difficulty in dealing with passive sentences as compared with actives. It was further predicted that they would have equal difficulty with all the passive sentences since they would not be able to use either the semantic or syntactic cues available in non-reversible and agent-deleted passives. For children at the next stage of development it was predicted that overall performance on passives would still be inferior to that on actives but that there would be differences in the pass rates for the different types of passives. The reversible passives would elicit more failures than non-reversible and agent-stated more failures than agent-deleted. Following from this it was predicted that passives

Table 1. *Sentence types*

Descriptive letter	Sentence type	Example
A	Active reversible	The dog licks the lamb
B	Truncated active reversible	The dog licks
C	Active non-reversible	The farmer rides the cow
D	Truncated active non-reversible	The farmer rides
E	Passive reversible	The lamb is licked by the dog
F	Truncated passive reversible	The lamb is licked
G	Passive non-reversible	The cow is ridden by the farmer
H	Truncated passive non-reversible	The cow is ridden

where neither syntactic nor semantic cues were present, namely those which were both full and reversible, would cause the greatest number of failures. No corresponding predictions were made about performance on active sentences since it was assumed that subjects would not need cues to interpret them. For the oldest children it was predicted that there would be no difference in performance on active and passive sentences since comprehension of the latter would be fully developed. Thus there would be no difference in this group between performance on reversible and non-reversible and agent-stated and agent-deleted passives.

METHOD

Design

The combination of each level of voice, reversibility and truncation resulted in eight types of sentence (A–H in Table 1). Four lists each containing 12 active sentences were drawn up. Two lists contained only reversible sentences (type A) and two only non-reversible (type C). Subtests B, E and F were constructed by rewriting each sentence in the reversible list as a truncated active, a passive and a truncated passive. Subtests D, G and H were constructed in similar fashion from the non-reversible sentences. The resulting eight active subtests (two of each type) were combined into two sequences within which each subtest (A–D) was derived from a different original sentence list. A further two sequences were derived by constructing the alternative versions of the reversible sentences (subtests A and B) by interchanging the subject and object. An analogous procedure was carried out to construct four sequences of passive subtests (E–H). One sequence of active subtests and one of passive was randomly allocated to each subject such that each sequence was used an approximately equal number of times in the experiment. This served to ensure that the semantic content of each subtest was balanced across subjects.

Pairs of line drawings were drawn for each picture in which the actor and object were either as described in the sentence or the opposite way round. For non-reversible sentences the incorrect picture depicted an anomalous situation.

The four subtests with active sentences were presented first in one session and the passive subtests in a further session. Within each session the order in which subtests were presented was randomized as was the order of sentences within each subtest.

Subjects

Fifty-eight children whose native language was English and whose ages ranged from 3:1 to 7:0 were selected from an infants' school, a playgroup and two nursery schools in central London. No child had a verbal IQ of less than 75 and 52 subjects had scores between 80 and 120. For the purposes of analysis subjects were divided according to mental age into three groups; those below 4:5 ($n = 10$, chronological age 3:1–5:0), those between 4:6 and 5:11 ($n = 21$, 3:10–6:9) and those above 6:0 ($n = 27$, 5:4–7:0).

Table 2. *Percentage of subjects passing each subtest*

Mental age (years:months)	*n*	Sentence type A	B	C	D	E	F	G	H
6:0–9:9	27	96	96	100	100	93	100	100	100
4:6–5:11	21	81	81	100	95	24	52	62	76
2:10–4:5	10	50	30	60	70	10	0	10	40
Overall	58	83	79	93	93	53	66	71	81

Procedure

Each child was presented with eight subtests, each containing 12 items of the same sentence type (A–H). Because of the length of the experimental task each child was seen individually for three sessions of approximately 15 min. The first session was devoted to the administration of the English Picture Vocabulary Test (Brimer & Dunn, 1973). Two or three days later the active sentences were presented. As these sentences were expected to be easier than the passives they were presented first to create confidence for the final test session, a week later, on the passive sentences.

Subjects sat next to the experimenter behind a table upon which each pair of pictures was placed in turn (left–right positioning being random). The instructions given to subjects at the beginning of the second session were as follows: 'Today I am going to show you some more pictures. This time you will have to look at two pictures and I want you to look carefully at both pictures and listen very carefully to what I say. I am going to tell you about what the people or animals or things in *one* of the pictures are doing. What I want you to do is to show me which picture I am talking about.'

Three pairs of practice pictures were presented which had one common element and one uncommon, for example, a lion chasing a bird/a lion climbing a tree. The instructions continued: 'Now then, you can see that these two pictures are not the same. Which is . . .? (e.g. The lion is chasing the bird).'

After the practice sentences the instructions were: 'The pictures I am going to show you now look *nearly* the same as one another but only *one* is the one I am talking about and I want you to show me which one it is.'

The experimenter then went through the four active subtests reading out each sentence in turn. Repetition and paraphrasing of instructions was sometimes necessary for the youngest subjects.

The procedure in the final session with the four passive subtests was identical except for the practice sentences which were all of type E (Table 1). During both the second and third sessions any spontaneous descriptions of the pictures by the subjects were recorded as well as their correct or incorrect choices of the pictures described by each sentence.

RESULTS

The subjects' correct or incorrect choices for each subtest of 12 sentences (maximum score = 12) were converted into pass–fail data by counting each score of nine or more correct choices as a pass and each score of eight or less as a failure. The percentage of subjects passing each subtest in each mental age group is shown in Table 2. The point of division had previously been determined by selection of a probability of 0·07 from the binomial distribution for $n = 12$ in preference to the stricter criterion provided by allowing only two errors for a pass performance ($P = 0{\cdot}02$). The choice of a less conservative probability was made in order to make some allowance for the inherently short attention span of young children. Some support for this decision was provided by a comparison between the pass–fail data resulting from the 0·02 and 0·07 division points which revealed that an equal number of active and passive subtest performances were rendered failures by the use of the lower probability.

Analysis of the pass–fail data for each age group using the Cochran Q test revealed that there was no significant difference in performance across subtests among the oldest children who had attained a mental age of at least six years ($Q = 8{\cdot}6$, $P > 0{\cdot}2$). In contrast, both the intermediate children (mental age 4:6–5:11) and those in the youngest group (mental age below 4:6) responded differentially to the subtests ($Q = 46{\cdot}6$, $P < 0{\cdot}001$; $Q = 24{\cdot}5$, $P < 0{\cdot}001$, respectively).

No further analysis of the oldest children's data was made in the light of their homogeneous performance across subtests. Comparisons of the subtest pass rates for the intermediate group, using sign tests, supported the predictions that non-reversible passives would be easier to comprehend than reversible ($P = 0{\cdot}003$)* and agent-deleted passives easier than agent-stated ($P = 0{\cdot}019$). Analogous comparisons of active subtest pass rates revealed that whereas non-reversibility also aided comprehension of actives ($P = 0{\cdot}02$) truncation did not ($P > 0{\cdot}05$). However, none of these comparisons involved all the members of the intermediate group ($n = 21$) since tied scores were dropped from the analysis. Twelve subjects passed all four active subtests so that tied scores reflected perfect performances. In contrast, tied scores in the passive subtest comparisons resulted mainly from unsuccessful performances, usually one subtest failure under each condition (e.g. reversibility vs. non-reversibility). Performance on active subtests was, however, superior to that on passive subtests even when tied scores of this kind were counted against the hypothesis ($P = 0{\cdot}013$).

In the youngest group ($n = 10$) five subjects failed all four passive subtests and of these, three also failed all the active subtests, selecting pictures at random. For the group as a whole, however, performance on active subtests was significantly better than on passive subtests ($P = 0{\cdot}008$). Further comparisons were not possible because of the small number of critical pairs. Certain general characteristics of the performance of the youngest group are discussed below.

The most interesting results are provided by an examination of changes in the use of syntactic and semantic cues with age. The performance on passive subtests may conveniently be divided into three categories:

CS	Complete success on all four passive subtests (including those with no cues, i.e. type E).
CCU	Consistent cue use, i.e. success only when either one or both of the experimental cues of non-reversibility or agent-deletion were available (failure on subtest E only or on E and F or on E and G).
NCCU	No consistent cue use, i.e. performance did not suggest that either cue had been consistently utilized, consistency being interpreted as success on both subtests where a particular cue was available.

The trend revealed in Table 3 of increasing success with mental age, initially with the aid of cues and eventually independently of them, is significant beyond the 0·0001 level ($S = 903$, $Z = 7{\cdot}19$).

Performances which fell into the CCU category were typically characterized by

---

* All significance levels are one-tailed.

Table 3. *Number of subjects falling in each performance category for passive sentences*

Mental age (years:months)	*n*	CS	CCU	NCCU
6:0–9:9	27	25	2	0
4:6–5:11	21	2	13	6
2:10–4:5	10	0	0	10

Note: For performance category notation see text.

failure on subtest E only. Indeed, of the 32 subjects who failed at least one subtest, 27 failed subtest E (84 per cent) compared with 20 subjects (63 per cent) who failed subtest F which caused the second largest number of failures. The proportion of subtest E failures is significantly different from chance using the binomial test ($P < 0{\cdot}0003$). The number of failures in conditions where one cue was present were very similar (subtest G caused 53 per cent of the 32 subjects to fail) but when both cues were present, in subtest H, only 11 subjects failed, representing only 34 per cent of those subjects who failed at least one subtest.

Children in the intermediate group who failed subtest E fell into three categories: those who used both the syntactic and the semantic cue and thus failed subtest E only ($n = 6$); those who used only the semantic cue and failed both reversible passive subtests ($n = 3$); those who used only the syntactic cue and failed both agent-stated passive subtests ($n = 2$). No subjects in the intermediate group succeeded only when both cues were present but three subjects in the youngest group behaved in this way, passing only subtest H.

DISCUSSION

The data presented here suggest that children who have attained a mental age of six years are able to comprehend passive sentences which contain neither cues of non-reversibility nor agent-deletion to aid interpretation. They are able to identify and correctly interpret the passive form solely on the basis of the differences in surface structure between active and passive sentences. Children below a mental age of six years experience greater difficulty with passive sentences, in contrast to children in the older group who comprehended all types of sentence with equal ease. However, children in the intermediate group (mental age 4:6–5:11) are able to utilize the cues to correct interpretation of passive sentences provided by non-reversibility and agent-deletion. The difference between reversible and non-reversible passive sentence comprehension is well established and no further discussion will be presented here. It is more surprising that there was a similar reversibility effect for active sentences in the intermediate and youngest group and this will be discussed below.

The most important finding of the present experiment is the confirmation of the prediction that agent-deletion would also facilitate passive sentence comprehension. As predicted, agent-deleted passives were easier to interpret than agent-stated passives whereas there was no such difference with active sentences. Although part of the explanation for this must lie in the greater familiarity of the agent-deleted form of the passive to young children, both in their own speech and in that of adults, another important factor may be that this form of the passive is not so likely to be converted into an actor–action–object active, since it only contains one noun, and thus is unlikely to be treated as 'typical' active. Spontaneous remarks made by subjects during the experiment give support to the idea that children found NVN actives 'normal' since the youngest children, particularly, tended to complete the truncated actives with an appropriate noun, for example:

E The farmer rides.
S The horse.

Spontaneous 'completion' of passives where the agent was not stated was far less common. Agent-deleted passives may thus be recognized as different from the majority of sentences with the accompanying realization that the NVN strategy is inappropriate.

Since both non-reversibility and agent-deletion appear to aid interpretation of passive sentences the particular difficulty experienced by subjects when either of these cues was present (i.e. in subtest E) is predictable. However, the pattern of performance on the other passive subtests suggests that the manner in which these cues are used is far from homogeneous. This is reflected in the comparatively wide mental age span of the intermediate group. Children of mental age around four years six months or less generally do not understand even the least confusable passives, that is, those which are both non-reversible and agent-deleted. The transition between complete inability to understand passive sentences and the stage at which reversible agent-stated passives are understood occurs mainly between mental age 4:6 and 5:11 and will probably be aided by an awareness of one or both of the syntactic and semantic cues investigated here. However, it is important to stress the large individual differences in the age at which passive sentence comprehension reaches completion. One subject as young as three years one month (mental age 4:9) passed all eight subtests whereas another child of similar mental age but who was chronologically two years older failed all passive subtests.

Although the youngest children performed badly on the passive subtests (see Table 2) there was some indication that they had begun to use cues. Four out of five subjects who responded differentially to active reversible and non-reversible subtests were more successful on the non-reversible sentences and a similar pattern existed for reversible and non-reversible passives. However, although four out of five passed more agent-deleted than agent-stated passive subtests, there was no such trend for truncated and full active sentences. The use of cues by the youngest subjects thus echoed the pattern of cue use in the intermediate group.

Before the data from this experiment had been collected there did not appear to be any reason to predict that children of three to six years would have difficulty in comprehending active sentences. Observations made during testing suggest that the (chronologically) youngest subjects had an attention span which was not only temporally short but also narrow. These subjects tended to respond to the first noun in the sentence (which without inspection of the verb did not distinguish the pictures) or to the first member of the picture pair which attracted their attention. This may partly have been due to prior experience with a vocabulary test which involved concentration on single words. However, general performance limitations of this kind, while accounting for an overall improvement in performance with age, cannot account from the significant difference between active and passive subtest pass rates in the two younger groups. The experimental technique would thus appear to have successfully tapped differences in voice comprehension over and above general task difficulties.

In the light of Bever's (1970) finding that children of around four years of age consistently use an inappropriate strategy for a short period, even to interpret nonreversible passives which younger children understand correctly, the small number of instances of consistently incorrect picture selection is perhaps surprising. Of the 75 passive subtest failures only eight resulted from consistent use of an inappropriate strategy (i.e. a score of 3 or less was obtained) and four of these cases occurred in one subject. One explanation of the failure to replicate Bever's finding may lie in the use of 12 instances to investigate comprehension. Whereas a child may treat a small number of passive sentences as though they were actives, the presentation of as many as 12 instances may give the child greater opportunity to become aware that the sentences being presented are *not* actives. When the child realizes this, he selects pictures at random because he does not know how to interpret the sentences. It was noted that some children did begin a subtest by consistently interpreting passive sentences as though they were actives but that on presentation of the fifth or sixth sentence their behaviour became random. Not all subjects who failed passive subtests behaved in this way, however, and it is likely that no single explanation of the failure to replicate Bever's (1970) findings is possible. The heterogeneous pattern of cue use in the intermediate group suggests that the processes involved in comprehending the passive voice will not be the same for all children.

This study, which was supported by a grant from the Social Science Research Council, was carried out under the supervision of Dr J. M. Greene in partial fulfilment of the requirements for the University of London PhD degree. The author is greatly indebted to Dr Greene for her comments and advice during the preparation of this paper.

REFERENCES

BEVER, T. G. (1970). The congitive basis for linguistic structures. In J. R. Hayes (ed.), *Cognition and the Development of Language*. New York: Wiley.

BRIMER, A. & DUNN, L. (1973). *English Picture Vocabulary Test Full Range Edition*. Bristol: Educational Evaluation Enterprises.

CROMER, R. F. (1970). 'Children are nice to understand'; Surface structure cues to the recovery of deep structure. *Br. J. Psychol.* **61**, 397–408.

FRASER, C., BELLUGI, U. & BROWN, R. (1963). Control of grammar in imitation, comprehension and production. *J. verb. Learn. verb. Behav.* **2**, 121–135.

GAER, E. P. (1969). Children's understanding and production of sentences. *J. verb. Learn. verb. Behav.* **8**, 289–294.

HERRIOT, P. (1969). The comprehension of active and passive sentences as a function of pragmatic expectations. *J. verb. Learn. verb. Behav.* **8**, 166–169.

MCNEILL, D. (1970). The development of language. In P. H. Mussen (ed.), *Carmichael's Manual of Child Psychology*, vol. I. New York: Wiley.

MARATSOS, M. P. (1974). Children who get worse at understanding the passive: a replication of Bever. *J. Psycholing. Res.* **3**, 65–74.

MENYUK, P. (1971). *The Acquisition and Development of Language*. Englewood Cliffs, N.J.: Prentice-Hall.

SCHOLNICK, E. K. & ADAMS, M. J. (1973). Relationships between language and cognitive skills: passive voice comprehension, backward repetition and matrix permutation. *Child Dev.* **44**, 741–746.

Sinclair, A., Sinclair, H. & De Marcellus, O. (1971). Young children's comprehension and production of passive sentences. *Archs Psychol.* **41**, 1–22.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *J. verb. Learn. verb. Behav.* **5**, 219–227.

Strohner, H. & Nelson, K. E. (1974). The young child's development of sentence comprehension: influence of event probability, nonverbal context, syntactic form and strategies. *Child Dev.* **45**, 567–576.

Svartvik, J. (1966). *On Voice in the English Verb.* The Hague: Mouton.

Turner, E. A. & Rommetveit, R. (1967). The acquisition of sentence voice and reversibility. *Child Dev.* **38**, 649–660.

Turner, E. A. & Rommetveit, R. (1968). Focus of attention in recall of active and passive sentences. *J. verb. Learn. verb. Behav.* **7**, 543–548.

(*Manuscript received* 16 *April* 1975; *revised manuscript received* 20 *August* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 429–438
Printed in Great Britain

STIMULUS PROCESSING AND RECOGNITION MEMORY IN CHILDREN

By DONALD H. SYKES
The Queen's University of Belfast

The recognition memory of eight-, ten-, and twelve-year-old children ($n = 108$) was tested using an incidental learning technique. The children were required to categorize a set of word stimuli according to either a structural (presence or absence of the letter A in the word) or a semantic (identification of the word as the name of an animal or a country) criterion. Recognition memory strength was similar for the three age groups for the semantically categorized stimuli. There was, however, a highly significant increase in memory strength with age for the structurally analysed stimuli. The results were interpreted as indicating that age-related differences in memory are a function of age-related changes in the development of stimulus-processing mechanisms.

In a number of recent articles on the development of memory in children (Corsini, 1971; Flavell, 1971; Jenkins 1971), the several authors arrived at remarkably similar conclusions. Essentially, these authors argued that memory is not a process or product which is separate from other psychological processes and products. It is a part of perception; it is a part of the cognitive activity of the individual, and as such its characteristics will reflect the characteristics, developmental and individual, of the perceptual and cognitive activities of the child.

This conception of memory has recently been presented in the non-developmental, experimental literature (Massaro, 1970; Craik & Lockhart, 1972). Briefly, Craik & Lockhart argue that existing experimental data concerning trace strength and the form or code in which the trace is held, are best understood in terms of the level or depth of processing to which the stimulus material was subject. Thus, material immediately present at the senses is subject only to a shallow level of analysis, an analysis which is largely concerned with the external physical and sensory features of the stimulus. Later stages in the perceptual/cognitive analysis of the stimulus will be concerned with the extraction of meaning from the stimulus, that is, with the semantic content of the stimulus as against its purely physical content. Finally, once the stimulus has been identified, there may occur a stage of stimulus elaboration and enrichment, when stored associations to the recognized stimulus are evoked.

There is then a continuum of perceptual analysis, ranging from a low level of sensory analysis to a deep level of semantic-associative analysis. Craik & Lockhart suggest that the by-product of this analysis is the memory trace, with the deeper analyses producing more persistent traces. In addition, of course, the level of analysis will determine the form or code of the memory trace. If the perceptual analysis does not proceed beyond the sensory stage, then the memory trace itself will reflect the physical characteristics of the stimulus. As the processing of the stimulus proceeds, so the code of the stimulus trace will change.

Irrespective of whether or not this hypothesis is correct, it has heuristic value, since it determines the questions that we might ask with respect to the development of memory in children. Within the perceptual processing framework, developmental questions would largely be concerned with developmental changes in the manner in which children process information and the consequence these changes have for memory. This latter approach has the additional virtue that it places the child's memory behaviour within the context of the child's perceptual and cognitive activities.

As Craik & Lockhart point out, one of the most readily available techniques for studying the relationship between the manner in which the individual processes stimulus material and his later memory for that material, is the incidental learning technique. The essential characteristic of this technique is that the subject processes the material in a manner which is compatible with the requirements of the task. By changing the requirements of the task, one can change the manner in which the subject processes the task material. By testing the subject's memory for the material involved in the task, one can then determine the relative memory consequences of the different processing activities. And this, essentially, was the experimental paradigm used in the present study, the particular technique being taken from a study by Schulman (1971).

More particularly, the study was designed to manipulate the depth to which the stimuli would be analysed during their initial presentation. Schulman (1971), with young adults, had found that an incidental learning task requiring the analysis of a stimulus in terms of its semantic content (its identity as a living thing or a geographical location) produced more accurate later recognition than a task requiring an analysis of the stimulus in terms of its structure (the presence or absence of the letter A). Similarly, studies by Hyde & Jenkins (1969), Johnston & Jenkins (1971) and Walsh & Jenkins (1973), among others, have demonstrated that in young adults semantic orienting tasks facilitate recall while non-semantic tasks do not. Developmentally, one might predict that following a semantic analysis of a set of stimuli, children of different ages, insofar as they are equally able to engage in this type of analysis, would demonstrate a similar degree of memory strength for those items. Furthermore, the memory strength consequent upon a semantic analysis should be greater than that following a structural analysis of the stimuli.

However, in the case of children, account must be taken of the degree of development of the component skills brought to the incidental task, skills concerned with the extraction and processing of information. Schaeffer (1973) has argued, in the context of cognitive development in children, that the skills younger children bring to a particular task, being less automated than is the case with older children, take up in their execution more of the limited processing capacity of working memory. In a structured task, which emphasizes speed of response, this relative lack of automatization of the appropriate skills might restrict the range of activities carried out on the stimuli, with negative consequences for later memory. With these considerations in mind, an additional monitor of the child's perceptual/cognitive analysis of the stimulus during the incidental learning task was taken, namely the latency from the presentation of the stimulus to the time that the child indicated his response. Craik & Lockhart (1972) have suggested that the time taken to process a stimulus may reflect the depth to which the stimulus has been analysed; the longer the latency the deeper the level of analysis. In the case of children, however, longer latencies cannot be assumed to reflect deeper levels of analysis. Younger children generally have longer latencies than older children across a variety of tasks and situations (Hohle, 1967; Wickens, 1974), the exact meaning of which is by no means clear. Wickens (1974) has suggested that, in addition to such factors as attentiveness, incentive and motivation, which affect reaction time, there may be some 'irreducible central-processing limitation difference relating to age'. It is possible that in children, particularly in younger children, longer latencies, rather than indicating a greater depth of processing, reflect a lesser degree of automatization of the component processes involved in the extraction and processing of information and the production of the required response.

METHOD

Subjects

In all a total of 108 children were tested, 36 children in each of the three age groups, 8 years ($\bar{X}$ = 8 yr 5 mth; S.D. = 6 mth), 10 years ($\bar{X}$ = 10yr 5 mth; S.D. = 3 mth) and 12 yr ($\bar{X}$ = 12 yr 5 mth; S.D. = 3 mth). Of the 36 children in each age group, one-half were boys and one-half girls. The children were selected from two schools, a primary and a secondary school, the secondary school being the continuation school for the majority of the primary school children. The schools were situated in a quiet, semi-rural residential area six miles outside a large industrial city. All children in the required age range were tested, the teachers being requested to exclude only those children who, in their opinion, were below their age level in reading ability. Each child was given the Vocabulary subtest from the WISC (Wechsler, 1949). Analysis of variance of the scaled scores with respect to age, sex and task condition to which each subject was assigned yielded only one significant factor, that being the sex factor, the boys having a significantly higher scaled score than the girls ($F = 4{\cdot}54$, d.f. = 1, 72, $P < 0{\cdot}05$: $\bar{X}$ males = 9·93; $\bar{X}$ females = 9·00).

Task and procedure

The task used in the study comprised two parts. During the first part of the task each child (tested individually) was required to categorize a set of 28 items, specifying in the case of each item whether or not it belonged to a particular category of stimuli. Upon completion of this part of the task, the child was then unexpectedly presented with a set of 56 items, and asked to state for each item whether or not it was one of the 28 items that he had just categorized. The 56 items included the 28 items that had been on the categorization list plus 28 new control items.

For the categorization part of the task there were six conditions, to which six children (3 male and 3 female) at each age level were randomly assigned. Children in condition A were instructed to identify in the set of 28 words those words that contained a letter A. Condition B was the reverse of condition A, in that the target words were words *not* containing a letter A. In condition C, all the words were target words, in that the subjects had to specify for each word whether it did or did not contain a letter A. The conditions A, B and C have in common the fact that correct categorization of each item requires that the subject merely note the presence or absence (depending upon the particular instructions given) of the letter A. These three conditions can be labelled the 'structural categorization conditions'. Conditions D, E and F are labelled 'semantic categorization', for successful categorization required the subjects to identify correctly the meaning of the words. In condition D, the target words were the names of animals. In condition E, the target words were the names of countries. In condition F, subjects were required to categorize each word as either the name of an animal or the name of a country, all items in this condition being targets.

For all of the categorization conditions, the general experimental instructions stressed that the children should work as quickly and accurately as possible. Recognition memory testing followed immediately after all 28 items had been categorized.

All 28 items in the categorization task were the same for all subjects in all conditions, and included the names of 14 animals and 14 countries, one-half of the animal and country names contining a letter A, the other half not. Similarly, the 28 control items were matched to the 28 stimulus items in these respects.

Measures and analyses

For the categorization task, the major dependent measure was the proportion of items correctly categorized. In addition the response latency, measured with a Forth Instruments millisecond timer manually operated by the experimenter, from presentation of the stimulus to the time that the child indicated his category response, was recorded.

For the recognition task there were two measures, the proportion of items seen before correctly stated to have been seen before (Hit Rate) and the proportion of items not seen before incorrectly stated to have been seen before (False Alarm Rate). These two measures were used to calculate the Signal Detection Theory measure of memory strength, d′ (Lockhart & Murdock, 1970; Banks, 1970). However, this measure requires stringent assumptions to be made concerning the underlying sensory distributions (McNicol, 1972), which assumptions are not always satisfied. Accordingly, as a check on the d′ measure of memory strength, certain non-parametric measures of memory strength were calculated, namely P*H (Snodgrass, Volvovitz & Walfish, 1972), C (Hammerton & Altham, 1971) and A′ (Grier, 1971). These three non-parametric measures, plus the parametric d′ measure, and the Hit Rate, were analysed separately. In all important aspects they produced essentially the same results. For the sake of brevity, only one measure will be reported here, namely d′.

The design of the present study was therefore a 3 (age) × 2 (sex) × 2 (type of categorization: structural vs. semantic) factorial, with separate analyses of variance being used to analyse the various dependent measures. In addition, separate analyses were carried out on the structural and semantic categorization conditions with respect to the type of categorization required. Thus, data from the structural condition were analysed using a 3 (age) × 2 (sex) × 3 (type of analysis: identify A words; non-A words; both types) analysis of variance, as were the data from the semantic categorization condition, the three levels of the third factor being, identify animals, identify countries, identify both.

Since many of the proportions in the categorization task approached 1·00, an arcsin transformation was applied to the data. A log transformation was applied to the latency data. Since the F ratios obtained with the transformed data were only marginally different from those based on the untransformed data, it is the untransformed data that are reported.

RESULTS

Categorization

Age was found to affect significantly the proportion of items correctly categorized ($F = 5{\cdot}32$, d.f. $= 2, 96$, $P < 0{\cdot}01$: $\bar{X}$ 8 yr $= 0{\cdot}95$; $\bar{X}$ 10 yr $= 0{\cdot}98$; $\bar{X}$ 12 yr $= 0{\cdot}98$). Also, response latency for correctly categorized items decreased with increasing age ($F = 35{\cdot}66$, d.f. $= 2, 96$, $P < 0{\cdot}001$: $\bar{X}$ 8 yr $= 2{\cdot}07$; $\bar{X}$ 10 yr $= 1{\cdot}64$; $\bar{X}$ 12 yr $= 1{\cdot}34$), as did the variance of within-subject latencies ($F = 7{\cdot}25$, d.f. $= 2, 96$, $P < 0{\cdot}01$: $\bar{X}$ 8 yr $= 0{\cdot}55$; $\bar{X}$ 10 yr $= 0{\cdot}19$; $\bar{X}$ 12 yr $= 0{\cdot}13$). In effect the findings point to a generally more skilled categorization performance by the older children.

Those children asked to categorize the items structurally did so significantly more accurately than those asked to categorize them semantically ($F = 13{\cdot}99$, d.f. $= 1, 96$, $P < 0{\cdot}001$: $\bar{X}$ structural $= 0{\cdot}99$; $\bar{X}$ semantic $= 0{\cdot}96$), this being particularly true for the younger children whose accuracy during structural categorization was on a par with that of the older children, but significantly lower during semantic categorization (age × type of categorization: $F = 4{\cdot}30$, d.f. $= 2, 96$, $P < 0{\cdot}05$), the means for the 8- 10- and 12-year-olds for structural categorization being 0·98, 1·00 and 0·98 respectively and 0·92, 0·97 and 0·98 for semantic categorization. The generally poorer performance under semantic categorization was due to the fact that while the children had no difficulty in correctly categorizing the names of animals, they did have difficulty in correctly categorizing the names of countries ($F = 25{\cdot}72$, d.f. $= 1, 36$,

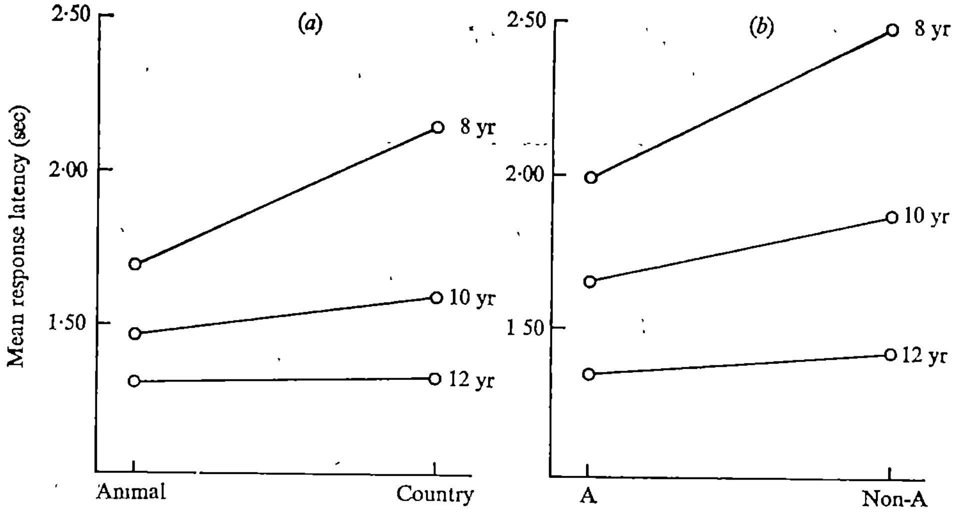


Fig. 1. Repsonse latency during categorization: (*a*) semantic categorization; (*b*) structural categorization.

$P < 0{\cdot}001$: $\bar{X}$ names of animals = 0·99; $\bar{X}$ names of countries = 0·92), this being the case at all ages, there being no significant interaction of age with this factor.

The response latency data indicated a significantly longer latency associated with a structural as compared with a semantic analysis ($F = 8{\cdot}74$, d.f. = 1, 96, $P < 0{\cdot}01$: $\bar{X}$ structural = 1·79; $\bar{X}$ semantic = 1·58). Furthermore, within the semantically defined conditions, the latencies to names of countries correctly categorized were significantly longer than the latencies to names of animals correctly categorized ($F = 11{\cdot}09$, d.f. = 1, 36, $P < 0{\cdot}01$: $\bar{X}$ names of animals = 1·48; $\bar{X}$ names of countries = 1·68), at least for the younger children (word type × age: $F = 4{\cdot}85$, d.f. = 2, 36, $P < 0{\cdot}05$; Fig. 1*a*). In addition, for all ages the mean latencies for the names of countries incorrectly categorized were significantly longer than the mean latencies for the names of countries correctly categorized (overall *t* test for correlated means = 6·19, d.f. = 31, $P < 0{\cdot}005$: $\bar{X}$ incorrect names of countries = 2·51; $\bar{X}$ correct names of countries = 1·77). It would seem reasonable to conclude that the children, particularly the younger ones, have less familiarity with the names of countries as compared to the names of animals, with the consequence that there is a greater degree of uncertainty in their behaviour when categorizing them, resulting in longer latencies and more errors.

Within the structural categorization conditions, the identification and categorization of a word as *not* containing the letter A was significantly more accurate than the identification and categorization of a word as containing a letter A ($F = 10{\cdot}24$, d.f. = 1, 36, $P < 0{\cdot}01$: $\bar{X}$ words without A = 1·00; $\bar{X}$ words with A = 0·98). There was no interaction of age with this factor. Words not containing the letter A, as well as being the more accurately categorized, also had the longer latencies ($F = 35{\cdot}97$, d.f. = 1, 36, $P < 0{\cdot}001$: $\bar{X}$ non-A words = 1·92; $\bar{X}$ A words = 1·66). At least, this was the case for the younger children, but not the 12-year-olds (word type × age: $F = 7{\cdot}47$, d.f. = 2, 36, $P < 0{\cdot}01$; Fig. 1*b*). It would appear that, in the case of the younger children, a negative identification (identification of word as *not* containing a

Table 1. *Mean d′, hit rate and false alarm rate for three age groups across two conditions**

	Categorization condition					
	Structural			Semantic		
Age (years)	Hit rate	False alarm rate	d′	Hit rate	False alarm rate	d′
8	0·43	0·20	0·97	0·68	0·11	2·19
10	0·62	0·12	1·78	0·76	0·07	2·44
12	0·66	0·11	1·98	0·63	0·03	2·34

* It should be noted that the reported d′ values are the group averages of d′ calculated on individual subject Hit rates and False alarm rates.

particular letter) takes longer than a positive identification. In addition, negative scanning of a word appears to place greater constraint on the children's behaviour, since it produces fewer errors.

Recognition memory

The manner in which items were processed during the categorization task had a significant influence on later memory for these items. Memory strength (as measured by d′) was significantly higher for the groups who had semantically, as opposed to structurally, categorized the words ($F = 30\cdot63$, d.f. = 1, 96, $P < 0\cdot001$: $\bar{X}$ semantic = 2·32; $\bar{X}$ structural = 1·58).

Memory strength was also related to age, but in a more complex fashion. There was a significant overall effect for age, indicating that memory strength increased with increasing age ($F = 7\cdot65$, d.f. = 2, 96, $P < 0\cdot01$: $\bar{X}$ 8yr = 1·58; $\bar{X}$ 10 yr = 2·11; $\bar{X}$ 12 yr = 2·16). However, this finding has to be qualified, for there was an interaction between age and the categorization condition. This interaction can best be understood when the structural and semantic conditions are analysed separately. With structural categorization, there was a highly significant age factor ($F = 11\cdot37$, d.f. = 2, 48, $P < 0\cdot001$), indicating that memory strength increased with age ($\bar{X}$ 8 yr = 0·97; $\bar{X}$ 10 yr = 1·78; $\bar{X}$ 12 yr = 1·98). In contrast, with semantic categorization there was no age effect ($F = 0\cdot54$, d.f. = 2, 48, n.s.), the three age groups having essentially similar memory scores ($\bar{X}$ 8 yr = 2·19; $\bar{X}$ 10 yr = 2·44; $\bar{X}$ 12 yr = 2·34).

The results also indicate that while memory strength is not as strong after structural as compared to semantic categorization, the difference in strength between the two decreases with increasing age (type of categorization × age: $F = 3\cdot52$, d.f. = 2, 96, $P < 0\cdot05$). This can best be seen in Table 1 where the d′ scores are presented, together with mean Hit Rates and False Alarm Rates. Further analyses indicated that, for all age groups, following semantic categorization the names of animals were better recognized than the names of countries ($F = 9\cdot82$, d.f. = 1, 36, $P < 0\cdot01$: $\bar{X}$ animals = 2·61; $\bar{X}$ countries = 2·28). There was no difference in memory strength following structural categorization for words with or without a letter A ($F = 0\cdot64$, d.f. = 1, 36, n.s.: $\bar{X}$ A words = 1·69; $\bar{X}$ non-A words = 1·60). However, the categorization strategy required of the subjects in the structural categorization conditions did have

Table 2. *Mean d′ scores across age and identification strategy (structural condition)*

	Identification strategy			
Age (years)	Positive identification	Difference score	Negative identification	Difference score
8	1·47		0·49	
		0·57		1·22
10	2·04		1·71	
		−0·24		0·87
12	1·80		2·58	

a significant influence on the later recognition performance of the younger children. If conditions A and B are compared, where condition A requires a positive identification (identify words *with* a letter A) and condition B a negative identification (identify words *without* a letter A), and condition C is excluded, since all words here were target words, there is a significant interaction of condition (A and B) with age ($F = 6{\cdot}03$, d.f. $= 2, 24$, $P < 0{\cdot}01$). This shows that a negative identification strategy results in relatively poorer memory performance in younger children (Table 2).

Finally, when memory strength was examined for targets across all conditions, except for conditions C and F where all the items were targets, the specification of an item as a target or a non-target item was found to have no effect on later recognition memory for that item.

DISCUSSION

The more accurate recognition of semantically categorized items relative to structurally categorized ones supports the interpretation that recognition memory reflects, in part at least, the stimulus processing or encoding strategy employed by the subject. Items analysed to a semantic level were more accurately recognized than items receiving merely a structural analysis. This finding agrees with that found by Schulman (1971) with young adults, and is compatible with the views of such writers as Massaro (1970) and Craik & Lockhart (1972), that memory is related to the perceptual/cognitive activity of the individual.

More informative, from the developmental point of view, was the finding that the age effects in recognition memory found in the study were also related to the encoding strategies employed. The younger children were significantly inferior to the older children in recognizing items from the structurally categorized set. These age differences cannot be ascribed to any hypothetically weaker memory stores or systems in younger children. If this were so, then one would expect poorer recognition memory in the younger children for the semantic items also. And this was not the case. Rather, it would appear that, when required to do so by the nature of the task, as in the semantic categorization conditions, the younger children process and extract as much information from the stimuli as the older children, in so far as their later memory strength for these items is similar to that of the older children. However, if the nature of the task is such that, as in the structural categorization conditions, not all the information contained in the stimuli needs to be processed for accurate performance, the younger children, apparently, do not process as much of the information as the older children, since during recognition their memory strength for these items is not as strong as that of the older children.

Detailed examination of the relative memory strength for structurally versus semantically categorized items indicated that, for the older children, their recognition memory strength following structural categorization was not dissimilar to that following semantic categorization (1·98 vs. 2·34). This suggests that while accurate performance in the structural categorization conditions does not require the processing of aspects of the stimulus other than the presence or absence of the letter A, none the less, the older children do appear to process at least some of the semantic content of the words, which then improves their later recognizability. If this is so, then the relatively poorer performance of the younger children with the structurally categorized items implies that they, unlike the older children, do not analyse the semantic content of the stimulus at the same time that they are analysing its structure.

Possibly, as Schaeffer (1973) has argued in the context of cognitive development in children, the component skills (perception, recognition, identification, etc.) brought to the categorization task by the children are less well automated in the case of the younger children. Being less automated they would take up in their execution more of the limited capacity of working memory, thus effectively preventing in the structural conditions the simultaneous analysis of content and structure within the limited time allowed (speed of response together with accuracy were emphasized in the instructions to the categorization tasks).

It could be argued of course that it is not that the younger children cannot engage in several activities at the same time but rather that they do not choose to. For it has been shown that the relatively poorer memory performance found in young children with certain memory tasks is not necessarily due to an inability to engage in a particular (mnemonic) activity, but simply in a failure spontaneously to do so (Hagen 1971). However, there is in this study evidence for the former explanation, which evidence is derived from the latency data. In the semantic condition, which led to equal memory strength for all age groups, the younger children took significantly longer to categorize the stimuli. This suggests that while the younger children have the appropriate analysing skills available to them, they are less well automated and therefore take longer to implement. In the structural conditions, which led to poorer memory in the younger children, the response latencies during categorization were again significantly longer for the younger children. More particularly, the structural condition which produced the poorest memory strength in the younger children was the same condition that produced the longest latencies in these children, namely the condition requiring a negative identification. It would seem reasonable to conclude that the longer latencies of the younger children in the structural conditions reflect less well automated component skills which, together with the particular task instructions and the emphasis on speed of response, effectively restrict the range of analyses carried out on the stimuli. A consequence of this is that later recognition memory for these items is relatively poor.

Certain other findings in the study deserve comment. Words defined as targets were no better recognized than words not so defined. Of far more importance in determining recognition performance was the manner in which the stimuli were analysed (structural vs. semantic, positive vs. negative identification) and the nature of the stimuli themselves. Irrespective of whether the words were targets or non-targets, the names of animals were more accurately recognized than the names of countries. During categorization, it will be recalled, the children were more accurate identifying the names of animals than the names of countries, which suggests that they are more familiar with the former type of stimuli. This being the case, they probably have a more detailed knowledge of the characteristics of particular animals than they do of the characteristics of particular countries. Accordingly, during categorization, when a word is identified as the name of an animal, this identification will produce in the child a whole host of associations and elaborations related to that particular stimulus. These associations could then serve to provide additional discriminatory cues aiding the later recognition of that item. Countries, being less familiar to the children, would engender fewer discriminatory cues, later recognition suffering accordingly. Additionally, the names of animals may be more effective in evoking related imagery than the names of countries, thus facilitating later recognition (Paivio, 1971).

In conclusion, it can be said that an important determinant of recognition memory in children is the nature of the orienting activity produced by the child in response to the task demands. Clearly, the study of memory processes in children requires that a detailed analysis be made of the component skills children of different ages bring to a task. For it is probable that age-related changes in memory reflect, not hypothetical changes in the capacity or strength of separate memory stores, but rather age-related changes in the organization of component skills concerned with the extraction and analysis of information from stimuli.

The author wishes to acknowledge the cooperation of the Down County Education Committee, the principals of the Holywood Secondary School and Redburn Primary School, and the children who participated in the study.

REFERENCES

BANKS, W. P. (1970). Signal detection theory and human memory. *Psychol. Bull.* **74**, 81–99.

CORSINI, D. A. (1971). Memory: Interaction of stimulus and organismic factors. *Hum. Dev.* **14**, 227–235.

CRAIK, F. I. M. & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *J. verb. Learn. verb. Behav.* **11**, 671–684.

GRIER, J. B. (1971). Nonparametric indexes for sensitivity and bias: computing formulas. *Psychol. Bull.* **75**, 424–429.

FLAVELL, J. H. (1971). First discussant's comments: What is memory development the development of? *Hum. Dev.* **14**, 272–278.

HAGEN, J. W. (1971). Some thoughts on how children learn to remember. *Hum. Dev.* **14**, 262–271.

HAMMERTON, M. & ALTHAM, P. M. E. (1971). A non-parametric alternative to d′. *Nature, Lond.* **234**, 487–488.

HOHLE, R. H. (1967). Component process latencies in reaction times of children and adults. In L. S. Lipsitt & C. C. Spiker (eds), *Advances in Child Development and Behavior*, vol. 3, pp. 225–261. New York: Academic Press.

HYDE, T. S. & JENKINS, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *J. exp. Psychol.* **82**, 472–481.

JENKINS, J. J. (1971). Second discussant's comments: What's left to say? *Hum. Dev.* **14**, 279–286.

JOHNSTON, C. D. & JENKINS, J. J. (1971). Two more incidental tasks that differentially affect associative clustering in recall. *J. exp. Psychol.* **89**, 92–95.

LOCKHART, R. S. & MURDOCK, B. B. (1970). Memory and the theory of signal detection. *Psychol. Bull.* **74**, 100–109.

MASSARO, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychol. Rev.* **77**, 557–567.

McNICOL, D. (1972). *A Primer of Signal Detection Theory*. London: George Allen & Unwin.

PAIVIO, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.

SCHAEFFER, B. (1973). Skill integration during cognitive development. Paper read at Conference on Current Research in Aspects of Long-Term Memory. University of Dundee.

SCHULMAN, A. I. (1971). Recognition memory for targets from a scanned word list. *Br. J. Psychol.* **62**, 335–346.

SNODGRASS, J. G., VOLVOVITZ, R. & WALFISH, E. R. (1972). Recognition memory for words, pictures, and words + pictures. *Psychon. Sci.* **27**, 345–347.

WALSH, D. A. & JENKINS, J. J. (1973). Effects of orienting tasks on free recall in incidental learning: 'difficulty', 'effort', and 'process' explanations. *J. verb. Learn. verb. Behav.* **12**, 481–488.

WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.

WICKENS, C. D. (1974). Temporal limits of human information processing: a developmental study. *Psychol. Bull.* **81**, 739–755.

(*Manuscript received* 28 *January* 1975; *revised manuscript received* 28 *July* 1975)

Br. J. Psychol. (1976), 67, 3, *pp.* 439–455
Printed in Great Britain

BOOK REVIEWS

The Psychology of Animal Learning. By N. J. Mackintosh. London: Academic Press. 1974. Pp. xiv + 730. £8.00.

To say that Mackintosh has produced the most comprehensive and scholarly analysis of the process of learning is not a sufficient complement to the achievement represented by his recent book *The Psychology of Animal Learning.* Other books on the topic provide neither the depth nor the breadth of coverage that this fundamental area of experimental psychology demands. Mackintosh has remedied this serious deficiency by providing a thoughtful and provocative analysis of the major issues in learning.

The book consists of a brief introduction followed by nine lengthy chapters. The methods and theories of classical and instrumental conditioning are dealt with in the first four chapters. One chapter is devoted to the analysis of aversively motivated behaviour, one to contrast effects and one to extinction. The final two chapters deal with generalization and discrimination. Although Mackintosh's writing style is precise and readable, the organization of the material within chapters tends to disrupt the logical progression of the narrative. Each chapter is divided into between three and eight sections many of which have a number of subsections. Some of the subsections also have subsections and the numbering and lettering of the many components of each chapter can be a source of some confusion.

The categories into which Mackintosh fits the findings of the (over 1800) articles that comprise his reference list will be familiar to students of psychology who have been nurtured on the writings of Tolman, Hull, Guthrie and Pavlov. He reduces the phenomena of learning to terms of the S–R-thing language. In so doing he has succeeded admirably in imposing a potentially comprehensible order upon experimental data gathered by researchers using a variety of methods and working from very diverse sets of assumptions. The organization of the published work on learning is in itself a substantial contribution to psychology.

Mackintosh does a masterful job in examining the experimental phenomena of learning in relation to the theories that have evolved to explain them. His method is first of all to provide an operational analysis of a particular learning phenomenon in which the empirical dimensions of the process are fully described. Thus, for example, in describing the determinants of the partial reinforcement effect the influences of the following variables are fully analysed: percentage of reinforcement, N-length, and N–R transitions, patterning of N and R trials; effects of reinforced and non-reinforced placements in the goal box; partial delay of reinforcement; effects of size of reward; intertrial interval; number of acquisition trials; generalization between non-reinforcement and punishment; effect of drugs. The list of topics submitted to this type of exhaustive operational analysis is impressively long.

Following analysis of the experimental findings related to a particular learning process, Mackintosh then goes on to examine the theoretical implications of these findings. Although in the introduction he denies commitment to any particular theory, a strong inclination towards a Tolmanian interpretation of learning is evident throughout the book. Often the resolution of a theoretical dilemma in favour of an S–S, expectancy interpretation appears to arise spontaneously although at other times 'expectancy' emerges as a victor by default: by the inadequacy of alternative (often S–R) formulations in accounting for the intricacies of the experimental data.

The process of expectancy is an intuitively appealing and parsimonious explanatory device that may be employed to interpret many behavioural phenomena. But it is not shown to be experimentally falsifiable. Perhaps this is why Mackintosh's theoretical conclusions frequently appear to be rather weak. Thus in commenting upon the observation of Shapiro *et al.* (1966) that salivation often accompanies panel pushing in dogs reinforced with food, he suggests 'that a more natural interpretation...is...that the presentation of the CS led subjects to expect that the time for the delivery of food, and therefore the time for responding, had arrived' (p. 228). The central problem of how reinforcement in instrumental learning establishes response-controlling properties in discriminative stimuli is treated in similar terms. Thus discriminative stimuli 'do not instigate responding: as a consequence of past learning they act as goals for responding'

(p. 233). Animals learn that 'their behaviour has certain consequences' (p. 211) and their response rate increases because of 'increased anticipation of reinforcement' (p. 232). Expectancy notions are also central to his interpretations of avoidance learning, conditioned inhibition and frustration (p. 346).

Biologically-based approaches to the analysis of behaviour such as ethology have provided some compelling criticisms of the artificial and intuitive behavioural categories employed by experimental psychologists. These criticisms tend to be supported by an evaluation of the theoretical analyses presented in this book. The analytical framework provided by traditional learning theory is only marginally successful. It lacks internal consistency, reducing the phenomena of learning to terms ranging from operationally defined processes to concepts still firmly rooted in mentalistic philosophy. In view of this, the reader may regret Mackintosh's reluctance to specify the assumptions underlying his theoretical preferences and to provide explicit definitions of fundamental terms such as *stimulus*, *reponse*, *motivation*, etc. Perhaps greater use of a biological frame of reference and of relevant data from the areas of physiological and comparative psychology may enable learning theory to progress beyond the formulations proposed 40 years ago by Tolman and Hull.

However, Mackintosh does not purport to generate a new approach to the explanation of learning processes and it is patently unfair to burden him with the sins of earlier generations. As a compendious reference source, a distillation of conclusions from much effort in animal laboratories, and as an example of methodology and the weighing of evidence, the book is a striking and welcome achievement.

JASPER BRENER and STEPHEN WALKER

Research in Psychophysiology. By P. H. VENABLES and M. J. CHRISTIE. London: Wiley. 1975. Pp. ix + 444. £12.00.

Venables and Christie have invited 17 contributions with which to interest the general or experimental psychologist, as well as the committed psychophysiologist, in the approaches of contemporary psychophysiology. Therefore this is not a specialist book, such as the substantive handbook edited by Greenfield and Sternbach, but rather aims to whet the reader's appetite with issues that may be pursued in depth elsewhere.

It is divided into three parts. The first concerns methodology. Lykken in a short chapter stresses the necessity of controlling for individual variability in group designs, and admonishes researchers for failing to look at score distributions in order to consider what may underlie variation from the mean. Mefford describes problems associated with statistical evaluation of longitudinal measures, and advantages in multivariate approaches. Christie and Todd point to ways in which psychophysiological measures may underline the influence upon the subject of the experimenter and test situation.

The second part is loosely called 'states'. It includes an excellent description by Frankenhaeuser of the behavioural correlates of sympathetic-adrenomedullary activity. Sayers introduces novel statistical analyses of sustained tonic changes in cardiovascular variables in laboratory and work situations. There are discussions by Johnson on the psychophysiology of sleep and by Naitoh on sleep deprivation. Other chapters include Bell's investigation of the menstrual cycle, and Christie's experiments relating skin potential levels to electrolyte balance and their variation with stress. Stern, Farr and Ray review studies of pleasure. The section ends with three chapters on psychiatric conditions: affective disorders by Lader and Noble, schizophrenia by Venables, and psychopathy by Hare.

In the final section physiological variables are shown to shed light on information processing. Sutton and Tueting tantalize with their studies showing the sensitivity of the cortical evoked potential in reflecting subtle attentional and motivational changes. Martin outlines the contribution made by autonomic conditioning to learning theory, drawing on her own meticulous work. Craik and Blankstein review the controversial literature concerning aspects of the psychophysiology of human memory. Finally, Venables, in pointing to the possible psychological and psychopathological implications of the recovery time of the skin conductance response, illustrates the advantage in interdisciplinary collaboration.

The editors' enthusiasm for psychophysiology is shared by this reviewer, who hopes that the collection awakens many experimental psychologists to the advantages in concurrent measurement of biological variables, and to methodological advances that psychophysiology has fostered. The reviewer has only a few reservations about the collection. The book is at its best when authors with established reputations in the field have been asked to write about their own research innovations. Considering the wealth of material to choose from, and remembering that the book aims to be representative of contemporary approaches, it is disappointing when a topic has been selected, about which psychophysiologists have made only scattered contributions, even though their potential to contribute is great. There is evidence too of an apparent rush to meet publication deadlines – typing errors, misspellings, chapters without adequate or any summary statements, and those in which there is repetitive reference to the next section.

Publication is nevertheless timely. Readers of this Journal may well wonder what the British Psychophysiology Group, now in its fourth year, organizes two-day meetings about. Wonder will increase when it is found that biofeedback has been deliberately excluded, and those critics who regard psychophysiologists as obsessed with technology will be heartened to find that techniques of measurement are no longer a major preoccupation.

J. H. GRUZELIER

The Psychology of Sex Differences. By ELEANOR E. MACCOBY and CAROL N. JACKLIN. Stanford, Calif.: Stanford University Press, and London: Oxford University Press. 1975. Pp. 634. £7.00.

Psychologists pay lip service to the notion of sex differences in intellectual and cognitive functioning, but rarely do they trouble to investigate such differences seriously. The result is the perpetuation of myths more befitting a psychological dark age than a time in social and economic history when women's roles and attitudes to women are being reassessed. Very early in training, the psychologist learns to 'control' for possible sex differences in experiments; before long this precaution becomes so ingrained that consideration of such differences as potentially important features of the behaviour under scrutiny is almost unthinkable. Thus sex differences are relegated to a position somewhere between psychological chickenfeed and scientific nuisance. Negative results are rarely published, and (in the interests of economy of space) positive ones are often omitted as well, since they are usually not embraced by the central hypotheses being tested. The consequent difficulty in evaluating this whole subject makes Maccoby and Jacklin's detailed and comprehensive review especially important, but one cannot help wondering how far the final picture represents the poor status of sex differences in psychological research to date and the lack of published data, rather than any underlying reality. This doubt is reinforced by the authors' tendency to talk in terms of what 'the majority of studies' in any particular area of investigation has shown, when that 'majority' is possibly a minority in real terms.

Despite these problems, this book is still an excellent work of reference for both researcher and student, and by today's standards, very good value for money. Slightly more than one-third of the volume consists of an annotated bibliography of over 1400 items, primarily covering research studies published since 1965. There are also 86 tables summarizing research findings in just about every area of psychological sex differences imaginable, from tactile sensitivity to parental pressure for achievement. Although a sequel to *The Development of Sex Differences* edited by Maccoby and published in 1966, except in the chapter on intellectual functioning, this new book is not intended to supersede that volume as the discussions of theoretical issues by various contributors contained in the earlier book, together with the annotated bibliography up to 1965, are still relevant.

The book does not attempt to deal with sexual behaviour as such, but rather is concerned with intellectual performance and social behaviours that are not specifically sexual but have been thought to be differentiated by sex. Throughout, the emphasis is upon *development*, and the final part of the book is devoted to a discussion of what the authors regard as the three main theories which purport to explain the development of such differences: observation learning, reinforcement learning, and self-socialization. There is no general overview of biological factors, but theories of hormonal, genetic, or other biological determination are discussed briefly in the text, particularly in relation to intellectual functioning, spatial ability, and aggression.

In considering possible genetic influences on sex differences, the authors concentrate on reviewing evidence for the sex-linkage theory, which relates spatial ability to a recessive gene on the X-chromosome. However, it is not pointed out that the sex difference in spatial ability has been reported for other organisms (e.g. rats) and has been shown to be affected by the sex hormones. Apart from an account of Broverman's activation theory, there is relatively little discussion of the role of sex hormones, or of their effect on embryological development. The sex chromosomes get a brief mention, but there is no discussion of their possible influence. Some consideration of Ounsted and Taylor's speculations on the '*Y*-chromosome message' would have been valuable here, as would an examination of sex chromosome abnormalities and their possible relevance to an understanding of the biological factors involved in sex differences. The overall treatment of biological matters is rather superficial, and gives the reader the impression that sex differences are *either* biological *or* environmental. Towards the end of the book there is a somewhat inadequate attempt to discuss the interaction of biological and environmental factors, but one is left with the feeling that the nature–nurture corpse has been clumsily resurrected.

Although there is continuing debate on the biological basis of sex differences in verbal, spatial, and numerical ability, and in aggressive behaviour, the existence of such differences is not seriously disputed. However, Maccoby and Jacklin refute the empirical basis for the female social stereotype of conformity, compliance, nurturance, dependence, passivity, and orientation towards social stimuli which many authorities have maintained represent real biological differences with evolutionary utility. Whilst it is true that some studies have shown females to be higher on these traits than males, such findings appear to be age-specific and the overall evidence does not support the hypothesis of such traits comprising a general personality cluster in females. However, where differences in these social characteristics have been found, it is very difficult to rule out the possibility of the stereotype influencing the expectations and judgements of the observers; this has been shown to be particularly likely when teachers' ratings are used, but it is doubtful if even the most naive observer can totally isolate himself from his cultural background.

The weight of theoretical discussion in this book is on the social origins of psychological sex differences. Traditionally, differential socialization of boys and girls has been regarded as a major factor in this, but although there is evidence of parental encouragement for children to develop specifically sex-typed interests and behaviours, and indications that boys generally have more intense socialization experiences than girls, in most other respects research suggests 'a remarkable degree of uniformity in the socialization of the two sexes' (p. 348). However, the inescapable fact that data are commonly derived from interview or questionnaire studies employing mothers (rarely fathers) and middle-class (more often than working-class) subjects must necessarily qualify any conclusion drawn in this area. The detailed investigation of child-rearing practices is still very much in its infancy and it is highly probable that the subtleties of differential socialization of the sexes are outside the scope of current techniques.

It is generally recognized, however, that reinforcement learning cannot by itself account for the acquisition of sex-typed behaviours, and therefore it is argued that imitation must play a significant part in the process. That observational learning occurs is not in doubt: the problem is why boys and girls should learn *different* things. The view that differences arise largely through selective imitation of same-sex models is disputed by Maccoby and Jacklin; they suggest that the evidence calls for some type of self-socialization process such as that proposed by Kohlberg. In this view, the child does not simply imitate, but rather attempts to match his own behaviour to his personal (and frequently very limited) conceptions of what maleness and femaleness are, and what males and females do and do not do, induced from what he has observed. This theory is particularly attractive as it does not treat sex-typing as special or even essentially different from the learning of other social behaviours such as language. The child is regarded as being faced with the task of operating effectively in a social environment, and he copes by hypothesizing the rules of the game on the basis of his accumulated experience and current cognitive sophistication, and then playing accordingly.

In some quarters it has been argued that culture and societies cannot *create* psychological differences between the sexes, they can only reflect and modulate those which are already biologically given. Maccoby and Jacklin contend that only in aggression and visuo-spatial ability does the evidence support such a notion. Otherwise, they suggest that this view is sterile, inaccurate, and socially iniquitous. They conclude wth the comments that societies have the option of minimizing, rather than maximizing, sex differences through socialization practices and that social institutions and social practices are not merely reflexions of the biologically inevitable. 'A variety of social institutions are viable within the framework set by biology. It is up to human beings to select those that foster the life styles they most value' (p. 374). For a book which has as its central thesis the social origins of psychological sex differences, it is particularly sad that the vast majority of research is based on subjects from white, middle-class, American culture. Arguably, in the future we would do better to look to cross-cultural studies to provide the test of this thesis.

C. H. SINGLETON and D. F. SEWELL

Mental Deficiency: The Changing Outlook. Edited by ANN M. CLARKE and A. D. B. CLARKE. London: Methuen. 1974. 3rd edition. Pp. xx+886. £16.50.

Since the mid-sixties, we have joined the Common Market, have gone over to decimal coinage, and inflation has accelerated. Since 1965, 'Clarke and Clarke' has lengthened itself by half again, has more than doubled its number of contributors, and has more than trebled its price: the rate of change of the changing outlook has quickened. Just as importantly, the book's structure has been altered from one based on a division between theory and practice, to a division by research area. Despite such fairly radical changes, the book's general message remains the same, '...that careful empirical work in all disciplines pays off both in the short- and long-term'. A British message, so general that it might be read into many handbooks; here, its fresh justification is found at a number of points, some of which are noticed in this positively prejudiced review. Structure first, then contents.

Earlier editions had three sections – *Background*, *Theoretical problems*, and *Practical problems*. The new edition has five – *I – Biosocial factors*, *II – The experimental analysis of subnormal behaviour*, *III – Assessment*, *IV – Amelioration*, and *V – Services* – titles that sometimes match their proper research areas with difficulty. For example, a difficulty of over-inclusion: Part I is concerned with classification, aetiology, epidemiology, the concept of intelligence, cognitive development, adoption and fostering, and longitudinal studies – matters pretty heterogeneous but easily comprehended within the part's title, *Biosocial factors* – as would be most of the rest of the book. Or, another difficulty: the title of Part II, *The experimental analysis of subnormal behaviour*, insofar as it might refer to a research area, would produce expectations about operant analyses, probably of the less complex forms of behaviour. It turns out to consist of three chapters, all valuable in their own right, but not mainly Skinnerian. One, by the editors, reviews a great deal of laboratory work – including some studies of operant learning – while the other two, by O'Connor and Hermelin, are concerned either with specific deficits or with complex behaviours like rule-learning and coding. Yet another difficulty is one of under-inclusion: a particular chapter - Mittler's on 'Language and communication', Tizard's on 'Longitudinal studies' - so much transcends its allotted place in the book as to challenge the editors' structure. But 'Clarke and Clarke' is now a grand design – and grand designs must have flaws.

What's in the book? In their Preface, the editors argue that there have been immense advances in research on mental retardation during the last decade. The same might be said of many other fields in behavioural studies but in the case of mental retardation there is one difference; that the starting point was much nearer to zero. For example, it was possible for the late D. Walton to publish, around 1960, a plea for the end of mutual and uncritical 'kindness' among researchers in mental deficiency – they should no longer have the status of Dr Johnson's woman preacher, but should be subject to the ordinary canons of criticism and reference to systematic evidence. And so they have been; advances have been immense. To this psychologist – and despite the editors' reference to 'all disciplines', Clarke and Clarke is far more psychological than anything else – the two most notable advances have been in the systematic attention given to the study of language, and in the development of behaviour modification. Both are treated handsomely here. Mittler's comprehensive chapter on language and communications is supplemented by a useful one on disorders of communication by Fawcus and Fawcus as well as by frequent references in other parts of the book. On behaviour modification, Kiernan makes a clearly written contribution – in Part IV, *Amelioration* – of a largely descriptive and illustrative kind, which must be by now a standard source for practitioners.

Part III, *Assessment*, three chapters long, is far more comprehensive than its predecessors. The first, by the editors, discusses capacity and performance in severe subnormality. In doing so, they point to a powerful source of possible conflict in matters of education and training; the opposition between the largely structuralist view, implicit in much progressive education, social work practice, Piagetian and Chomskian psychology, that, as they put it 'the best learning is self-generated' and the more functionalist view, supported by recent empirical work and implicit in Gunzburg's later chapters on education and training, that in teaching mentally handicapped people a prime requirement is for very well-structured learning situations. D. F. Clark, in the second and third chapters on assessment, addresses himself mainly, and usefully, to specific techniques but does not omit to recognize that assessment takes place in a social context, about which the assessor needs to have a systematic awareness. Clark's discussion of measures of social competence is a sparse by comparison with the fulness of the rest of his two chapters.

What's not in the book? Neither a full biology, nor a sociology, of mental retardation. Berg has revised his concise chapter on pathological factors in aetiology, to which is added a new chapter by Kirman, on individual differences having a known genetic or pathological basis. Clarke and Clarke have never attempted an exhaustive coverage of biomedical aspects, for which fuller treatments can be found in other texts; but they do claim both a psychological and a social emphasis, and it is now at least arguable that their treatment of the social is one-sided. Despite three chapters of high quality – Kushlick and Blunden's on epidemiology, Carr's on families, and Tizard's on services – there is nowhere represented in the book a particularly sociological point of view. Edgerton's work on stigmatization is recognized, but it says something for the recency of much sociological interest in mental retardation that later American work on labelling is not mentioned, presumably because it emerged too late for the current edition.

So, on to the fourth edition. With it in mind, Clarke and Clarke are adjured to crack their editorial whip more sharply at their contributors, to renounce confusing use of the phrase 'scientific concept', and to publish simultaneously in soft and hardback – say, £350 and £500, around 1980?

JOHN CASTELL

Aberrant Development in Infancy. Edited by N. R. ELLIS. New York: Halsted Press (Wiley). 1975. Pp. viii+287. £8.85.

This collection of papers, although covering a wide range of topics, centres around a common theme of how and why normal development fails to occur in some human infants. All were prepared for and presented at the seventh annual Gatlinburg (Tennessee) Conference on Research and Theory in Mental Retardation.

Faced with a diverse collection of studies using both human and animal material, the 'human' prospective reader tends to look rapidly down the list of contents or to search hopefully for a foreword which will help him to select those topics of most interest to himself. However, his searching behaviour may be accompanied by an uneasy feeling that he is failing to grasp the flavour of the conference as a whole, and a most useful feature of this collection is the concluding chapters in which four independent overviews are made of the material, and in which some of the general themes and assumptions underlying the work are critically presented.

The papers fall largely into three categories: what the future holds when early development is aberrant; the implications of the use of pharmacological agents on mental retardation and other aspects of behaviour; and, finally, the assessment of early behaviour, especially in normal infants, as a predictor of later intelligence.

The first of these categories is headed by a paper in which Isaacson urges the reader to take a highly critical look at his ideas on recovery from early brain damage. Subsequent studies, which use largely animal subjects, range from that of Denenberg and DeSantis who used the effects of restricting the blood supply to foetal rats to provide animal models for the small-for-gestational-age infant, to Berkson's observations on the pattern of social responses of macaque monkeys caring for blind infants.

As might be expected, work on the effects of pharmacological agents also relies a good deal on animal material, but a study is reported by Kaplan *et al.* which was carried out in the context of a programme to refine diagnostic and treatment methods of neonates suffering from narcotic withdrawal. Tantalizing references are made to evidence presented by Brackbill, on the effects of commonly used analgesics and anaesthetics on the behaviour of offspring – tantalizing because although her work is presented as highly relevant by the 'overviewers', it is not presented in the volume itself.

Two closely related areas of direct concern to those working with infants in a clinical setting are the identification of infants 'at risk', and the early assessment of infant behaviour as a predictor of later intellectual development. The original methods of compiling risk registers frequently produced lists of such an unwieldy size that the amount of action which could be based on them was limited. Moreover, many infants in groups selected on the basis of a single factor do tend to develop normally. Parmelee *et al.*, looking ultimately towards helping the community to make the best use of its resources, present some initial data on their attempts to measure a wide range of variables and to use multiple assessments at different ages. They see the strength of their approach as lying in the possibility of identifying the contributions of their measures both independently and in combination so that eventually a more efficient system can be defined for identifying those infants for whom there is an increased probability of handicap in childhood.

The development of the 'visual interest' test by Fantz in 1956 provided a useful tool for many researchers looking at the behaviour of infants, and important contributions to this volume are three papers by Fagan, Cohen and Friedman which explore the idea that habituation may provide a basis for the assessment of intelligence. Using the fact that infants spend more time fixating a novel or non-habituated stimulus, a measure can be made of recognition memory. It is the individual differences among infants on this measure which is put forward as being potentially important for the prediction of later intellectual functioning.

One merit which cannot be denied to Piaget's theory is that it continues to prod us to try to understand infancy. A paper of great interest to those who share this view with its author is that by Kopp; she presents a classification of fine motor behaviours as observed in eight-month-old infants and relates the resulting data to the sensory–motor schemes used by her subjects. Again, looking ahead, this is another approach which may have implications for the prediction of future abilities.

A good deal of the work reported in this volume is in its early stages and represents only a part of studies which are being carried out on a much larger scale; the result is a collection of stimulating papers which leave the reader looking forward with interest to the outcome of the many projects proposed and also perhaps with a fresh approach to researches of his own.

JEAN ROBERTSON

Parent–Infant Interaction. Ciba Foundation Symposium 33 (new series). Amsterdam: Associated Scientific Publishers. 1975. Pp. viii+324. $21.95.

These proceedings of a symposium held in 1974 include 15 papers, edited discussions, and an introduction and summing-up by the Chairman (Hofer). There were 25 participants, mainly American and British, including psychologists, paediatricians and zoologists. Two papers are on rhesus monkeys, one is on rats and the rest almost entirely on human neonates.

There have been several symposia on mother–infant interaction and one wondered that the Ciba Foundation thought it necessary to organize another. It seems to have been worthwhile, for several reasons. Techniques and experimental design have become more sophisticated in a very short time, and so also have thinking about infancy and scepticism about previous assumptions; also there are several new and sometimes surprising findings (some experiments obviously need replication). The variety can be illustrated by highlights: observations on the transition from hormonal to non-hormonal control of maternal behaviour in the rat (Rosenblatt); detailed 'naturalistic' observations of mothers' behaviour with their infants immediately after birth, including cross-cultural differences, and emphasizing the contrast between the elated interactional behaviour of mothers who have home deliveries compared with those in hospital (Klaus *et al.*); the regular cycling of attention in the infant at three weeks when taking part in face-to-face games with the mother, and the dejected response when the mother ceases to attend, or if the mother is blind and cannot respond to the infant's repertoire of expressions (Brazelton *et al.*); evidence that olfactory discriminations have been learnt by eight days of age, in that babies orient more to their own mothers' breast pads (Macfarlane).

This last paper is one example of the very early learning which is frequently reported, and which illustrates a radical change in thinking about a neonate's abilities. Learning which a decade ago was thought to occur at so many months of age is now believed to occur at so many weeks, or even days. Another theme which appears in these papers is that the mother's behaviour may be much under the 'control' of the baby, and that for many developmental research purposes the mother–infant pair is the best unit to consider. Perhaps the most outstanding feature of this symposium was the lack of certainty about what predictions can be made for the future of the child. In one well-known longitudinal study (Dunn) it has proved impossible to make any important predictions from detailed study of mother–infant pairs, but in another study (Kennell and Leiderman) there were shown to be correlations between the amount of contact a mother had with the infant in the first few days and her behaviour to the infant months later. Participants seem to have become wary over what behaviour is 'good', partly because correlations are sometimes the reverse of what might be expected. For instance, a slow reaction to interruption of sucking in the neonate is related to a slow rate of smiling in infancy but to greater expression of pleasant emotions at pre-school ages (Bell). Authors generally showed a concern with the relevance of their work to human development (Oppé's paper is specifically on the relevance to paediatrics) but I think it would be fair to say that no clear predictions about future mental health, etc., can be made from the evidence presented here. Kennell and Leiderman showed that premature babies (and their families) are at risk if parents are not allowed to interact with them, but, as the Chairman pointed out, it is not clear how long lasting the effects may be. Dunn suggests that individual differences in early mothering and infant behaviour (which are related to variations in labour and delivery) may not be good predictors of later interaction because 'normal' mother–infant pairs (none of her sample is premature, etc.) show great resilience to and compensation for early inadequacies.

What then should neonatal care be like? Bowlby says: 'Although we are all interested in the long-term effects of particular practices, we know that (relevant) information is extremely difficult to get. My view is that we have to use our common sense by adopting criteria which are near at hand and which seem likely to be relevant.... A mother who is confident and enjoying what she is doing is almost certainly going to do far better for her baby, both physiologically and psychologically, than if she is anxious and miserable.' One concludes that the symposium is of considerable scientific rather than practical interest.

B. M. FOSS

Juvenile Theft: The Causal Factors. By W. A. BELSON. London: Harper & Row. 1975. Pp. xix + 411. £15.00.

Major investigations of juvenile delinquency are few and famous and Belson's book is likely to join them. Its aim is to examine causal factors in juvenile theft using a hypothetico-deductive approach, but this book could equally be evaluated as an informative text or reference book on juvenile theft.

The study avoids some of the more common difficulties of delinquency research. It deals only with theft, rather than with the heterogeneous behaviours which constitute delinquency, and 44 different types of theft are given separate consideration. It considers various criteria for measuring theft: value of the objects taken, variety of type of theft and frequency of theft. It attempts to get at the full range of theft and is not restricted to the behaviour of boys apprehended by the police; in fact a small deficit in the otherwise comprehensive coverage of the investigation is its exclusion of approved school boys.

There is a convincing and obsessional thoroughness about the approach to the study and a very adequate description of the procedures involved. Great care is taken to get an appropriate sample of boys and the noteworthy efforts to create the conditions under which boys may admit to theft have beeen described by Belson previously. However there is no validation of the self-report procedure and while one agrees with Belson that this is difficult, even small partial validations would have been useful.

The main content of the book is a mass of data on the stealing behaviour of 1425 13 to 16 year old boys based on a long interview with each, arranged in chapters with separate aetiological themes such as getting caught for stealing, association with thieves, parental control and punishment of boys, etc. As a source book of factors associated with juvenile theft, the book is unsurpassed.

However this is a lower aim than the causal one held by the author. Forty-five causal hypotheses were examined by deducing a series of expectations from each and comparing boys who qualified on the aetiological criterion with those who did not, using a complex matching procedure based on a composite of factors correlated with stealing. So for example, the hypothesis 'Association of a boy with thieves increases his own involvement in stealing' led to 12 testable expectations including 'The greater the proportion of a boy's associates who steal, the more likely is he himself to steal', 'The longer the period of a boy's association with thieves, the more likely is he himself to be involved in stealing', and 'Boys will claim that when they started to associate with thieves, their own stealing started or increased'. The main disadvantage of this approach is that it examines causal factors retrospectively and that most of the hypotheses not only originate in studies which were also retrospective but also probably arose from the type of retrospective material which is investigated in the testing of expectations. This is particularly critical since West (1973) has shown that it is possible to execute a prospective study of juvenile delinquency, although he does not examine the same range of hypotheses.

Belson's approach does, however, give a very comprehensive range of data and by considering what support there is for each hypothesis, the author makes a series of recommendations for remedial or preventive action. These could be useful in two ways: firstly as a possible set of targets in the rehabilitation of juvenile thieves and secondly as a series of possible experiments for the investigation of causal factors.

MARIE JOHNSTON

WEST, D. J. (1973). *Who Becomes Delinquent?* London: Heinemann.

Language in Childhood: Form and Function in Language Learning. By HAZEL FRANCIS. London: Paul Elek. 1975. Pp. viii + 240. £5.95.

The author, helpfully, indicates two principal aims of her book, to extend published empirical evidence and to discuss theories of language acquisition. The second of these is tackled mainly in the initial and final chapters, and the first aim in the intervening five chapters.

The empirical work is derived from Dr Francis' MA and PhD theses, and some of it will already be familiar to interested researchers. These chapters present limited evidence on a number of only partially overlapping topics. Chapter 2 discusses early speech and presents analyses of situational variations in the speech of one two and a half year old, finding rather more evidence of word-class variations than utterance-type variations. The following chapter continues with more detailed analyses, both syntactic and semantic (rather than functional, as the author claims), of the same child. These lead the author to stress the importance of cognitive and functional, rather than grammatical, factors in early language development. Chapter 4 reports data from groups of five to seven year olds on story telling and sentence repetition tasks, finds more signs of social class differences in vocabulary than syntax, and expresses scepticism regarding the extent to which a seven year old can be said to possess a strictly linguistic competence. Chapter 5 reports interview and discrimination task studies of the early stages of literacy of the same groups of children. These studies reveal some of the difficulties children experience in mastering written language and, interestingly, suggest that it is easier to detect the impact that the acquisition of written language has on the child's speech than to document relations between prior speech skills and subsequent reading and writing. The final empirical chapter presents tests of the seven year old's ability to operate in the abstract with language, e.g. to inflect nonsense forms in the absence of referential support, and again concludes that there is little evidence of a strictly linguistic competence viz. 'there is no strong reason to hypothesize any knowledge of a deep structure related to surface structure by transformation rules' (p. 188). Both details from Dr Francis' findings and some of her conclusions are obviously of interest, although, at some points, there seem hardly enough of the former quite to justify the latter.

The theoretical contributions, which impart some sense of unity to the book as a whole, are, in 1975, somewhat limited in their impact. Chapter 1 is a rather piecemeal and simplistic review of theoretical positions and methodological problems, and like the final chapter has a rather dated air. In this last chapter Dr Francis continues her criticism of the Chomsky–McNeill view of language acquisition in terms of an innate grammatical competence and sketches out her own alternative framework, which is a combination of elements of both pre- and post-Chomsky–McNeill views.

She accepts some role for traditional associative and reinforcement learning, and in addition stresses cognitive, semantic and functional components, which are very much in tune with many recent proposals from other writers but the author does not relate her suggestions in any detail to these developments. Had her discussion appeared in the early seventies, when one suspects it was originally written, it might well have aroused interest; now it seems merely to point to a number of important topics of current concern without getting very far with any of them.

Thus, Dr Francis' book is likely to prove of most interest to developmental psycholinguists unfamiliar with her research but working on topics on which she reports her own findings.

COLIN FRASER

Children's Spatial Development. Edited by J. ELIOT and N. J. SALKIND. Springfield, Ill.: Charles C. Thomas. 1975. Pp. ix+296. $23.50.

During the past decade there has been a considerable increase in interest in verbal and linguistic processes – psycholinguistics, acquisition and development of language, social class and language, teach-your-baby-to-read, and so on. In contrast, there has been less emphasis on spatial behaviour; the appearance of a book on children's spatial development is, therefore, a refreshing change.

The editors' purpose in compiling this volume was to stimulate research in a relatively unexplored area. In order to do this the book is divided into three sections: the first contains four essays, each written by a different researcher working on spatial problems; the second consists of résumés of 40 studies chosen to give a cross-section of the sort of work being done in the general area of children's spatial abilities; and the last presents an extensive bibliography covering a wide range of literature on spatial behaviour.

Although the first section does not claim to present a comprehensive overview of the field, the four essays considered together give reasonable coverage though none of the contributors deals with, for example, the young child's difficulty in representing another's view of objects in space.

The first essay, by Lauren J. Harris, concerns the neurophysiological factors underlying spatial skills; it makes fascinating reading. The evidence presented strongly suggests that, in the infant, the right hemisphere of the brain (the 'spatial' side) is more active than the left (the 'language' side); spatial skills develop more rapidly than linguistic ability, and do not depend on prior development of language skills. It follows from this that the early development of cognition might well be called the development of spatial knowledge – the child's developing knowledge of the spatial layout of his world. Harris's case is well argued and presented in a clear and easy-to-read style.

In contrast, the essay by S. G. Vandenberg concerning sources of variance in the performance of spatial tests is less satisfactory. This is largely because only two sources of variance are discussed: social class differences (rather briefly) and sex differences; variations resulting from differences in age, intelligence and ethnic group are not considered. An attempt is made, however, to clarify the meaning of *spatial* ability as opposed to any other; it is clear that different authors use the term 'spatial' in different ways. Vandenberg also draws attention to the need for spatial tests appropriate for young children; he considers that a lack of suitable tests presents a major obstacle to further research efforts.

David R. Olson's essay deals with the relationship, as opposed to the difference, between spatial and linguistic processes. Olson points out that some linguistic problems may be treated spatially, whereas some spatial problems may be treated linguistically; moreover, the same problem may be treated either spatially or linguistically. The process chosen to solve a particular problem will depend on the nature of the task, its complexity, and the subject's competence in translating the task into one symbolic system or the other. The development of the language of space and its relationship to spatial understanding is also considered; the evidence presented suggests that the structure or organization of our perception of space underlies or determines the structure of our language.

George E. Forman's essay attempts to draw a parallel between the historical development of scientific conceptions of space and its development in the child. Just as each new theory (from Aristotle to Einstein) represents successive decentrations from a geocentric view of space, so each stage of the child's thought represents a further decentration from his egocentrism. Forman's essay presents, as the author admits, a somewhat disjunctive comparison of the development of spatial concepts in child and man, and because of this it is rather difficult to follow.

The 40 abstracts presented in the middle section of the book represent summaries of a cross-section of studies concerned with spatial behaviour. On the whole, these are well chosen and give a fair indication of the range of current investigations. Although these abstracts are meant only to be a starting point for new researchers, some are so abbreviated that it is difficult to get much idea of what they are about; a good feature, however, is the list of further references suggested after many of the abstracts.

The most disappointing feature of the book is the last section, the bibliography. Like the middle section, this extended list of references is intended as a source for the intending researcher.

Amazingly, however, one of the most important works on children's spatial development is omitted, viz. Piaget and Inhelder's *The Child's Conception of Space*. Generally, the studies summarized in the middle section of the book are not listed in the bibliography; it would have been more helpful to have had all the references together so that the bibliography could be used independently. There are also many defects such as a lack of consistency in the use of different kinds of type face and in the format of the references. More seriously, some references are incomplete, names such as Michotte and Warburton are misspelt, and some of the references mentioned in the four essays are missing from the bibliography. Since the bibliography is such an important part of the book these faults are especially damaging because they undermine its main purpose: to provide an authoritative and up-to-date source book in this area of psychology. Despite its shortcomings, as the book is the first attempt to gather together such a wide range of material on children's spatial development, it may still fulfil its aim. MAUREEN V. COX

Experimental Psycholinguistics: An Introduction. By SAM GLUCKSBERG and JOSEPH H. DANKS. Lawrence Erlbaum. 1975. Pp. xiii+233. £5.40.

The psychology student wishing to study psycholinguistics has to suffer handicaps he would have avoided had he chosen a more traditional area, like visual perception or animal learning. First, since psycholinguistics is so eclectic a discipline, relevant findings and ideas are published in a much wider variety of organs, many of which will not be available in departmental libraries. Second, it is a discipline which undergoes frequent and dramatic reorientations, and thus makes it hard – especially for undergraduates and non-specialists – to forecast, even in general outline, the shape of future developments, and to pick out of the current mélange of studies those destined to be more than historical curiosities. Third, the student may have to put up with the absence of a specialized psycholinguistics course, or with a course devised and taught by a psychologist who has not specialized in it.

Therefore, the availability of a good introductory text is of particular importance to students. Unfortunately, no really satisfactory psycholinguistics text exists. Authors who have attempted it have, usually, sacrificed coverage for a coherent interpretative framework – which was possible, even sensible, during an age of faith. (More agnostic authors – like Herriot – offered the reader alternative approaches but no overall synthesis.) Glucksberg and Danks's efforts illustrate some of the difficulties of trying to write a textbook in a period peculiarly characterized by crumbling certainties and widening horizons (an age of reason?); for example, the groping attempts to locate language more firmly in its social matrix.

The student has a right to expect of an introductory text that it should provide enough coverage to get a good overview of existing work; that the topics covered are organized in such a way that, even if there is no integrating theory, clearly defined concepts will explicate the links among studies by various authors, or from different points of view, or from different disciplines; that there is a critical evaluation of past work, *and* an indication of the direction of future developments. It would also be nice if the text were written so as to communicate some of the excitement of the research and the ideas leading up to it.

If he judges by these criteria, the student will, I am afraid, be disappointed with this volume. First, coverage: encouragingly, the book contains discussions of recent material not found in previous introductory texts, for example, work on memory for prose by Bransford and his colleagues, a chapter on acoustics and phonology, a chapter on dialects and schooling (the treatment of which is, incidentally, polemical psycholinguistics, not experimental). There are, however, glaring omissions. The authors say in the Preface that the task of psycholinguistics is 'to discover how people produce and comprehend speech': only two pages are concerned with speech producing, and neither contains any reference to experimental studies. The student might well conclude that this task has been so totally neglected by psycholinguistics that only *a priori* speculation is at present possible. Another serious omission concerns the psychological work on the semantic and pragmatic functions of syntactic form. Well-known studies are not mentioned, including those by Wason, Greene, Johnson-Laird and Huttenlocher, whose significance has been continually reaffirmed by subsequent research. Yet the early experiments on transformations – which really are now just historical curios – rate a few paragraphs. And so the student could easily get the impression that variations of syntactic form are empty of any semantic or pragmatic significance.

However, the experiments that are covered are generally well described, and the student will appreciate the clear presentation of method and rationale, and the well-judged and pertinent critical evaluations – though, to my taste, the authors have been too sparing in their criticism of some of the experiments they discuss. The pity is that the book is just too fragmentary: the experimental findings, well treated in isolation, are not 'integrated...[into] an intelligible framework' as the authors promise in the Preface. For example, chapter 3 is called 'Word meanings and the mental dictionary' and treats semantic feature theory, word associations and Quillan type models of 'semantic memory', but does not show how they might be relevant to each other, or indeed, to chapter 4, which deals with sentence comprehension, and in spite of the fact that here there is a section on case grammar, which could, as in Lindsay and Norman's introductory text, be used to coordinate these areas of investigation. Likewise, chapter 6, on 'Dialects and schooling' mentions the cross-dialect imitation experiments of Baratz; clearly these have some bearing on sentence comprehension – but neither is mentioned in connexion with the other. Chapter 7, called 'Language, thought and communication', should deal with topics which other authors, by no means obscure, have related to understanding and remembering sentences, to dialect and school achievement, and to social rules for conversations; but, in fact, the chapter is restricted almost totally to studies of object- or colour-naming within and across languages.

This fragmentariness is, I think, a reflexion of the fact that the authors are primarily experimentalists, not theorists. But this does make it hard for the student who, as a novice, has to find the common ground among the chapters, particularly since the authors have played Humpty-Dumpty with some of the key linking terms – old favourites like 'grammar', 'meaning' and 'context'. 'Grammar' – which here is to include semantics and phonology too – is sometimes used to refer to these 'Operations...logically sufficient to produce speech' (p. 2) and other times to 'an explicit description of the language' (p. 4). 'Context', as is all too common in the literature, can refer to anything that is not in the acoustic signal, with the implication that all such things act in the same way. 'Meaning' is divided into two sorts: 'referential' and 'denotative', which, as I understand their account, distinguishes the actual object referred to by a word in an actual utterance, from the class of objects to which the word can refer, ignoring the constraints of a particular utterance context. Nothing turns on this distinction in subsequent discussion, except to explain why sentences like

Washington is the President
Washington is Washington
The President is the President

differ in meaning; but in this case the logically better established distinction of 'sense'/'reference' (Frege) would be appropriate.

Those with more than a passing acquaintance with psycholinguistics will be able to benefit from the many useful insights scattered through this volume, and indeed the rather haphazard arrangement and selection of topics may lead them along unfamiliar avenues and stimulate them to create new syntheses of their own. But the beginning student will need a text which can impose a little more order on the present chaos.

BRIAN BUTTERWORTH

Human Memory: Structures and Processes. By R. L. KLATZKY. San Francisco: W. H. Freeman. 1975. Pp. 276. £7.20 cloth; £3.90 paper.

This book conforms to the prevailing tendency to regard man as an information-processing system. The vicissitudes of stimulus input as it proceeds through the successive stages of the sensory stores, pattern recognition, short-term storage, and long-term storage are well documented and clearly described. Indeed, the strength of the book lies in the exceptional clarity with which we are given a conducted tour through well-charted territory. For example, two topics which many students find somewhat confusing are Sternbergery and the applications of signal-detection theory to memory. Klatzky provides the clearest exposition of both topics (and several others) that I have encountered. From the point of view of the undergraduate student, this exceptional limpidity of style has much to recommend it. However, as we shall see, there are two sides to this particular coin.

There are at least two main approaches that can be taken when writing about a body of experimental research: the descriptive and the evaluative. The descriptive approach involves spelling out in some detail the various characteristics of the relevant experimental paradigm and some of the main variables investigated. A comprehensive account of typical findings and a representative theoretical explanation of these findings is then given. Klatzky usually follows this approach. For example, when discussing work within the Sternberg paradigm, she spends several pages focusing on details of the paradigm and on presenting Sternberg's theoretical contributions. No other theories are considered, and the only suggestion that something may be seriously wrong with Sternberg's theory is contained in a footnote! To cite another example, Klatzky spends 13 pages describing signal-detection theory and its relevance to recognition memory, but scarcely refers to any problems. There is, for example, no mention of the difficulties occasioned by the frequent non-independence of the measures of sensitivity and criterion.

The evaluative approach involves a brief description of typical experimental findings, followed by a lengthier and critical examination of the major theoretical explanations. In essence, the descriptive approach reduces psychology to a small, formal garden with well-delineated paths, whereas the evaluative approach confronts the student with the inchoate reality of a weed-filled and relatively untended wilderness. The reviewer greatly favours the evaluative approach, since this facilitates the process of mental scythe development, a necessary component of the student's armamentarium.

An important omission is any serious consideration of recent theoretical emphases on word attributes and on levels of processing. In spite of the title of the book, Klatzky has organized her book in terms of ideas about the structure of human memory rather than in terms of processes. There is, of course, some mention of processing theories, but the work of Craik and Lockhart, of Cermak, of Hyde and Jenkins, and of Underwood on memorial attributes is not discussed at all.

A further important weakness is the failure to discuss what Reitman (1970) referred to as the 'decoupling' problem. He pointed out that psychologists frequently assume that one aspect of human functioning can be studied in isolation, an assumption that is surely erroneous. While Klatzky does attempt to show the interdependence of perception and learning, she does not consider either the relevance of motivational factors to learning and memory, or the involvement of problem-solving strategies in retrieval. Current experimental practice is certainly to try to decouple the memory system from other systems, but it is probable that the attempt is both unsuccessful and misguided.

In sum, the book provides a reasonably satisfactory introduction to the field of human memory. Most of the main areas of contemporary research interest are discussed competently and at some length. It seems likely, however, that students of an inquiring turn of mind would prefer a rather more challenging text.

REITMAN, W. (1970). What does it take to remember? In D. A. Norman (ed.), *Models of Human Memory*. London: Academic Press.

MICHAEL W. EYSENCK

Theories of Group Processes. Edited by CARY L. COOPER. London and New York: John Wiley. 1975. Pp. x+277. £7.95.

At first glance one's reaction to this volume is, 'Here, at last, is what we've been waiting for – the book that Lewin unfortunately did not live to write'. Disappointment, however, soon follows. The title is a misnomer. The book turns out to be somewhat less ambitious. But first a word of clarification about research on group processes. As those familiar with groups research will know, the field divides roughly into two areas: the *experimental* (group productivity, risky shift, etc.) and the *experiential* (effectiveness of T-group training, interpersonal trust in groups, etc.). The present volume is an assembly of papers mainly about the experiential area. And here it is necessary to fall back even further. The book is not an account of theories of experiential group psychology; it is, instead, an assortment of papers about theory-building in this area – something rather different. In the words of Professor Cooper himself, 'This volume will focus...on theory-building in the area of experiential groups and organizational processes'.

Having indicated the scope of the book, the ground is clear for a look at the contributory chapters. There are ten chapters in the book. The first, by Professor Cooper himself, is really a brief introduction to research in group psychology, and an outline of 'intent' for the present volume. It is stimulating and informative – Cooper is obviously someone who is very much at home in this rather uncertain field. For these reasons one regrets that he did not go on to develop his introduction into a full-length chapter. The following five chapters of the book are concerned with 'global' aspects of theory-building in experiential group psychology. Argyris, who leads off, explores the relationship between intrapersonal variables, group-learning environment variables and organizational/individual variables. An impressive opening – and very 'global'. Argyris is followed by Kolb and Fry who deal with applications of theoretical ideas in experiential learning; they also consider the functional importance of approach and style in this. Continuity of theme can be seen in the chapters up to the sixth by Blake and Mouton, in which the relationship between group-organization and the individual-in-the-group is explored by investigating the effects of different kinds of intervention.

From chapter 6 onwards, a complementary general framework emerges, in which change in the group (or in the individual-in-the-group) is examined against some particular group-variable. For instance, Golembiewski and McConkie examine the relevance of ***trust*** to ***group/individual change***, and Reddy deals with the relationship between ***group composition*** and ***individual change***.

While the two general frameworks described emerge from a perusal of the book, they do not come up readily. Indeed, one gets the impression at times that most of the contributors were quite unaware that any kind of framework was intended, and although Cooper draws attention to the two general frameworks which can be seen, he disclaims any intention of having made provision for them when planning the volume ('the intention was not to restrict the contributors to a set topic but to issue them with a licence to build theoretical schemas in those areas of group processes of their own choosing', p. 4). So, although it would be agreed that the book, as it stands, has a certain amount of structure, one feels that more would have been helpful – whether the book is regarded as a research manual or as a student text. Editorial linkage too, between the chapters, would also have been helpful. Here, however, there is some schematic aid in the form of Cooper's fig. 1.

In case the prospective reader of the volume is, by now, considering putting aside any immediate intention of reading the book, with an, 'Oh dear, another of them', one must quickly turn to the individual chapters and point out that some of these, as 'separates', are particularly enlightening about research work in this area and very readable. Reddy's chapter, for instance, in which interpersonal affection and change in sensitivity-training groups is discussed, is refreshingly succinct and comparatively free from jargon. Similarly, the chapter by Lakin and Constanzo on the leader and the experiential group, is a model of cogency with readability.

On the whole, one would say that the book is stimulating, if at times something of a 'rich mixture'. It sets a high standard for later additions to the 'Individuals, Groups and Organizations' series of which it is the first volume.

JAMES G. M'COMISKY

Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences. By J. Cohen and Patricia Cohen. Hillsdale, NJ: Lawrence Erlbaum. 1975. Pp. xxi + 490. £12.25.

The fact that regression and correlation models preceded analysis of variance models by several decades and were developed and used in a different context has obscured, until recently, their conceptual similarities, especially in non-mathematical textbooks available to behavioural scientists. The main aim of this book is not only to demonstrate their similarity, but also to show that both can be regarded as special cases of a wider analytic system. It is to this wider system (MRC) that the title of the book refers and, indeed, in the closing chapter, it is argued that MRC is itself a special case of the yet more general system of canonical correlation.

The text is intended both as an introduction for non-mathematical behavioural science students and as a working manual for research workers. Hence, as a text, it offers a verbal-intuitive exposition, rich in redundancy and concrete examples and, as a research manual, it offers an integrated conceptual system and practical working methods. This combination of structural approach with practical instruction, together with the conceptual clarification of links among different statistical methods, produces a clear and exciting text which represents a significant advance, both conceptually and pedagogically, in this notoriously difficult field of application of statistics to behavioural science.

The first step in the development of the wider conceptual system is achieved by a shift of emphasis in the use of multiple correlation and regression in behavioural science which, traditionally, has been for forecasting with only incidental attention to explanation. This book reverses the emphasis, its almost exclusive concern being with the analytic use of MRC for purposes of explanation. The main explanatory devices centre on extension of the familiar concepts of 'proportion of variance accounted for' and 'partialling'; these are obvious links, for example, between MRC and analysis of variance.

The first three chapters explain the basic structure of MRC and cover very much the same ground as would be included in the conventional text on multiple correlation and regression. It is in Chapters 4 to 10, the heart of the book, that the novel and elegant superstructure is added. This involves four main new constructions, viz. methods of representing *sets* of independent variables, the use of these sets as units of analysis, the concepts of partialling and of hierarchical models of independent variables and the generalization of interactions as conditional relationships. These innovations and developments produce a general system of data analysis which is not only applicable to most types of data, illuminating many research problems, e.g. missing data problems, in the process, but also opens up real possibilities for new methods of data analysis. Chapters 9 and 10, which demonstrate the MRC generalization of analysis of covariance, repeated measures and matched subjects experimental designs, are particularly valuable in this respect.

The final chapter argues that this general model is itself a special case of a generalized version of canonical correlation. Again, one of the key gains is the linking of correlational and experimental models, in this case multiple correlation and multivariate analysis of variance and covariance. Three useful mathematical appendices are given in addition to usual statistical tables. One offers a mathematical treatment of the basis of MRC, one offers desk calculator solutions and a third offers useful information for selecting appropriate computer programmes.

In total, an impressive book which breaks new ground, conceptually and pedagogically. For those wanting a simple and clear, yet rigorous, non-mathematical introduction to multiple correlation, multiple regression and analysis of variance, this book can be strongly recommended. For those interested in the conceptual structure and general development of data analytic models, it is a must.

Only two misgivings disturb this reviewer's enthusiastic appraisal of this pioneering text. Both stem from the authors' data analytic approach which considers only the question of 'given the data, what kind of analysis yields what kind of meaning?' Firstly, this approach blurs the important distinction between primary and secondary data and, secondly, it allows the use of variables of different and doubtful status of measurement, reliability and validity such that assumptions of the statistical model used may be violated. It assumes that the researcher will be sufficiently expert to interpret the data with due caution and in relation to the statistical assumptions of the methods used. This touching faith in the statistical wisdom of the majority of behavioural scientists seems hardly justified in the light of their current performance in published research.

A. B. ROYSE

Psychosocial Stress: Population, Environment and Quality of Life. By L. LEVI and L. ANDERSSON. New York: Spectrum Publications. 1975. Pp. xv + 142. $5.60.

It is difficult to know what to expect of a book with such an all-embracing title as this one has and yet contains fewer than 100 small pages of text. The table of contents confirms the impression that the authors are in danger of attempting to tackle an impossible task for a short book. The foreword by Hans Selye contains the pretentious claim that, 'This book will render an invaluable service to humanity by enriching the Quality of Life...and represents a great contribution to human welfare'. I trust the authors will be suitably embarrassed by this statement because their book does not merit such acclaim – but then how many books ever did?

In fact this volume was prepared at the request of the Swedish government 'to give an overview of existing research on human response to environmental pressure'. Still a tall order, but if a very limited definition of 'overview' is adopted the book comes much closer to meeting this claim than Selye's.

Levi and Andersson have picked up most of the fashionable concepts of ecological psychology and given them the very briefest of treatments backed up by an equally limited but sensible bibliography. In the first of the eight chapters they define the problem as they see it and certainly touch on important issues. When it is realized that they describe a theoretical model for the study of population, environment and quality of life in under two pages and devote one paragraph to the section headed 'stress and disease', the inevitable superficiality of the book becomes clear. They move on to present, in the second section of the book, some ideas on the effects of population density and movement, the urban environment, industrialism, unemployment, cultural and social factors, the physical environment, and nutrition on the quality of life. They also identify three groups who are at greatest risk of being unable to cope with the pressures of the environment, namely children, old people and the physically, mentally and socially handicapped. On Levi and Andersson's implicit definition these groups would seem to include everyone in the world except a few young adult males in North America and Western Europe. Neither at this point nor elsewhere in the book are the implications of the authors' enthusiastic categorizing and summarizing considered.

The book finishes by considering the interrelatedness of man and the total environment and suggesting an 'integrated world plan of action'. This laudable plan is presented in two pages and consists of an argument for more research and the monitoring of environmental pressures and a table of human needs. There are no practical suggestions.

The weakness of this book is not that it has glaring gaps but that the whole fabric is too thin to be of any use except to someone who has no prior contact with the vast amount of work done on human response to social stress and who is able to use the book as a source of some useful references rather than as a text in its own right. The authors make this kind of usage somewhat difficult by presenting the findings of selected studies without considering them in the context of the literature and so making it very easy to construe the results of one study as a summary of what is known about a particular topic.

RAYMOND COCHRANE

Unmasking the Face. By PAUL EKMAN and WALLACE V. FRIESEN. Englewood Cliffs, NJ: Prentice-Hall. Pp. 212. Price not given.

Paul Ekman and Wallace Friesen have carried out a lot of valuable research into different aspects of the facial expression of emotion. In a previous book they reviewed the evidence and concluded that six emotions are expressed in the face in all cultures studied so far – surprise, fear, anger, disgust, happiness and sadness. They compiled an 'Atlas' consisting of photographs of each part of the face for each emotion, and showed that the expressions from the Atlas corresponded with the subjective experience of emotion.

The present book is primarily a training manual based on this research, for learning to recognize emotions from facial expressions. There is a chapter on each of the six emotions, showing the respective shapes of forehead, eyebrows, eyes and eyelids, nose, cheeks and mouth. Anger for example involves three areas – brows, eyes and lips. Different variations are distinguished, for example four kinds of surprise; different degrees of intensity are shown, as well as blends between two emotions. Here it is admitted that the scheme goes beyond established research findings – which so far show that only six main states can be distinguished reliably. Research has not yet established in detail the information sent by different parts of the face for each emotion, or their relative importance. There is also the problem that different people express the same emotions somewhat differently. In addition to the training materials there is a conceptual analysis of the processes involved in expressing and not expressing emotions.

A training procedure is suggested, and 54 photographs of several performers are provided to be cut out and used as training materials. No evidence is provided that this training is effective, though I hope to find out myself in the near future. The information in the text, and the collection of photographs, go beyond anything previously available, and could well be the basis for developing more sensitive awareness to emotions shown in faces. On the other hand, in real life, facial expressions are often less clear than those shown in this book, which were produced by moving specific muscles under instruction – even though some involve only part of the face, and some are blends of more than one emotion; in everyday life we often inhibit our facial expression. However in real situations we have a lot of other evidence to go on – the nature of the situation itself, other non-verbal cues, especially tone of voice, and the contents of speech.

When people are given social skills training it is sometimes found that they need to be made more sensitive to the emotions expressed by others, or at any rate to be more attentive and observant. More important however is their own expression of emotions. A rather cumbersome procedure is suggested for 'checking your own facial expression', involving taking 14 photographs, though a simpler procedure with a mirror is mentioned as an alternative. The results show whether one is an ***unwitting expressor***, a ***blanked expressor***, a ***frozen-affect expressor***, an ***ever-ready expressor***, or a ***flooded-affect expressor***. It might be more useful to discover a person's characteristic expression, and his expressiveness in this channel.

This book should be of great interest to psychologists interested in emotions, faces or non-verbal communication, and to those involved in social skills training for whom it presents a unique set of training materials. Although it does go beyond the published evidence, it seems very likely that the main contents of the Atlas are mainly correct. Though the face is not the only signal for emotions, it is certainly one of the main ones.

MICHAEL ARGYLE

Problems of Philosophy and Psychology. By JAY N. EACKER. Chicago: Nelson-Hall. 1975. Pp. xii+201. $11.00.

It is a source of embarrassment to psychologists and delight to some philosophers that psychology finds it so difficult to shake off its philosophical origins. It is natural therefore that psychology undergraduates should make some contact with philosophy, as they are required to do in courses at some universities. These courses demand an elementary textbook. Professor Eacker has not succeeded in writing it.

Eacker's book is a survey of the literature on 15 philosophical problems of psychology. The sources he quotes are usually secondary, and consist of psychologists' accounts of the views of philosophers. Eacker does not pretend to offer solutions to the problems he discusses.

The best feature of the book is the nine-page glossary of philosophical terms which occurs at its end.

PETER GIBBINS

New Mind, New Body. By BARBARA B. BROWN. London: Hodder & Stoughton. 1975. Pp. xiii+464. £5.25.

The title of this book sets one's alarm bells ringing. Perplexed perusal of the list of contents does nothing to assuage these fears: 'Skin talk: conversations with the subconscious'; 'Heartbeats: the heart that nobody knows'; 'Alpha brain waves: the importance of being alpha'.

If one ignores these warnings, and I do not advise it, the book fully lives up to expectations. It is generally euphoric in its tone, expounding the present findings and future miracles of biofeedback with breathy enthusiasm, stops on the way to wage war on experimental psychology – altogether too systematic and dull for Dr Brown – and enjoys the discomfort that she wrongly imagines biofeedback causes Skinnerians and other opponents of the autonomous will. The book's size and recent publication date suggest it might be a useful source of summaries of the field: it is neither, being both very selective and dated.

Altogether, an irritating and unhelpful book, of little interest to the readers of this journal.

DEREK JOHNSTON

Br. J. Psychol. (1976), 67, 3, *pp.* 457–459
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

ADCOCK, C. & REICH, B. *Values, Attitudes and Behaviour Change.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

ALTNER, G. (ed.). *The Nature of Human Behaviour.* London: George Allen & Unwin. 1976. Pp. viii + 467. £7.50.

BARBER, P. J. & LEGGE, D. *Perception and Information.* London: Methuen. Pp. 144. £2.50; paper, 80p.

BARHAM, R. M. & BOERSMA, F. J. *Orienting Responses in a Selection of Cognitive Tasks.* Rotterdam: Rotterdam University Press. 1975. Pp. xiii + 154. D.fl. 39.00.

BARNETTE, W. L. (ed.). *Readings in Psychological Tests and Measurements.* (Third edition.) Baltimore: Williams & Wilkins. $7.95.

BECKER, E. *Escape from Evil.* London: Collier-Macmillan. 1976. Pp. xix + 188. £4.50.

BECKER, E. *The Denial of Death.* London: Collier-Macmillan. 1976. Pp. xiv + 315. £1.50.

BENDER, M. P. *Community Psychology.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

BERRY, J. W. & LONNER, W. J. (eds). *Applied Cross-Cultural Psychology.* Amsterdam: Swets & Zeitlinger. 1975. Pp. viii + 338. D.fl. 31.00.

BURGESS, T. *Home and School.* Harmondsworth: Penguin. 1976. Pp. viii + 246. 80p.

DAVIS, J. C. & FORSYT, J. P. (eds). *Mental Examiner's Source Book.* Springfield, Ill.: Thomas. 1975. Pp. x + 238. $14.50.

DAVIS, L. E. & CHERNS, A. B. *The Quality of Working Life.* Vol. 1. *Problems, Prospects and the State of the Art.* London: Collier-Macmillan. 1976. Pp. xiii + 450. £7.75; paper, £3.00.

DAVIS, L. E. & CHERNS, A. B. *The Quality of Working Life.* Vol. 2. *Cases and Commentary.* London: Collier-Macmillan. 1976. £7.75; paper, £3.00.

DE MAUSE, F. (ed.). *The History of Childhood.* London: Souvenir Press. 1976. Pp. ii + 450. £5.00; paper, £3.00.

DINITZ, S., DYNES, R. R. & CLARKE, A. C. *Deviance, Studies in Definition, Management and Treatment.* (Second edition.) London: Oxford University Press. 1976. Pp. xiii + 641. £5.95.

EYSENCK, H. J. & EYSENCK, S. B. G. *Eysenck Personality Questionnaire (Manual).* Sevenoaks: Hodder & Stoughton. 1975. £1.65.

FEATHER, N. *Values in Education and Society.* London: Collier-Macmillan. Pp. xv + 350. £9.95.

FREUD, A., EISSLER, R. S., KRIS, M. & SOLNIT, A. J. *The Psychoanalytic Study of the Child.* London: Hogarth Press. 1975. Pp. xiv + 780. £11.00.

GILES, H. & POWESLAND, P. F. (eds). *Speech Style and Social Evaluation.* London: Academic Press. 1975. Pp. viii + 218. £6.60.

GLEN, F. *The Social Psychology of Organizations.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

GOLDSTEIN, J. H. *Aggression and Crimes of Violence.* London: Oxford University Press. 1976. Pp. xiv + 192. £4.75.

GREEN, M. *Goodbye Father.* London: Routledge & Kegan Paul. 1976. Pp. viii + 169. £3.95.

HAWKES, L. C. *Pragmatics of Analoguing. Theory and Model Construction in Communication.* London: Addison Wesley. 1975. Pp. xv + 199. £3.00.

HEATHER, N. *Radical Perspectives in Psychology.* London: Methuen. 1976. Pp. 142. £2.50; paper, 80p.

Hinde, R. A. & Hinde, J. S. *Instinct and Intelligence.* London: Oxford University Press. 1975. Pp. 16. 45p.

Hyman, H. H., Wright, C. R. & Reed, J. S. *The Enduring Effects of Education.* Chicago: University of Chicago Press. 1975. Pp. ix + 313. £7.50.

Jonas, D. & Jonas, D. *Sex and Status.* London: Hodder & Stoughton. 1976. Pp. 214. £4.50.

Kahn, S. *Essays in Freudian Psychoanalysis.* New York: Philosophical Library. 1976. Pp. xx + 159. $10.00.

Kaplan, F. *Dickens and Mesmerism. The Hidden Springs of Fiction.* Princeton: Princeton University Press. 1975. Pp. xv + 250. $10.00.

Kaplan, M. F. & Schwartz, S. (eds). *Human Judgment and Decision Processes.* New York: Academic Press. 1976. Pp. iii + 325. £7.55.

Kessel, L. & Byers, P. *Observations on the Training of a Surgeon.* University of London Institute of Orthopaedics Symposium 2. London: Institute of Orthopaedics. 1976. Pp. 90. £1.50.

Kirby, R. & Radford, J. *Individual Differences.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

Krippner, S. & Rubin, D. (eds). *The Energies of Consciousness.* London: Gordon & Breach. 1976. Pp. xvi + 243. £5.60; paper, £2.50.

Kroy, M. *Moral Competence: An Application of Moral Logic to Rationalistic Psychology.* The Hague: Mouton. 1975. Pp. 356. D.fl. 75.00.

Lawson, R. B., Goldstein, S. G. & Musty, R. E. *Principles and Methods of Psychology with Laboratory Manual and Kit.* London: Oxford University Press. 1976. Pp. xvi + 528. £8.50.

Mackay, D. *Clinical Psychology: Theory and Therapy.* London: Methuen. 1976. Pp. 142. £2.50; paper, 80p.

Mercer, C. *Living in Cities. Psychology and the Urban Environment.* Harmondsworth: Penguin. 1976. Pp. 240. 75p.

Morris, K. I. & Cinnamon, K. M. (eds). *Controversial Issues in Human Relations Training Groups.* Springfield, Ill.: Thomas. 1975. Pp. xxvii + 137. $9.95; paper, $5.95.

Murrell, H. *Motivation at Work.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

Murrell, H. *Men and Machines.* London: Methuen. 1976. Pp. 144. £2.50; paper, 80p.

Pick, A. (ed.). *Minnesota Symposia on Child Psychology.* Minneapolis: University of Minnesota Press. 1975. Pp. viii + 204. $12.50.

Pickard, P. M. *If You Think Your Child is Gifted.* London: George Allen & Unwin. 1976. Pp. 160. £4.50; paper, £2.25.

Potts, J. *Beyond Initial Reading.* London: George Allen & Unwin. 1976. Pp. 115. £3.95; paper, £1.60.

Roazen, P. *Freud and his Followers.* London: Allen Lane. 1976. Pp. 613. £10.00.

Sack, O. W. *Awakenings.* Harmondsworth: Penguin. 1976. Pp. 344. £1.00.

Serpell, R. *Culture's Influence on Behaviour.* London: Methuen. Pp. 144. £2.50; paper, 80p.

Stacey, B. *Psychology and Social Behaviour.* London: Methuen. 1976. Pp. 142. £2.50; paper, 80p.

Stevenson, Y. *The Hot-House Plant. An Autobiography of a Young Girl.* London: Paul Elek. 1976. Pp. xvi + 190. £6.00; paper, £3.00.

Webb, J. *The 5 Ideologies of Progressive Education.* Chorley: Greensilver. Pp. 48. Price not given.

Weiner, M. L. *The Cognitive Unconscious. A Piagetian Approach to Psychiatry.* New York: Yale International Psychological Press. Pp. xix + 202. $12.95.

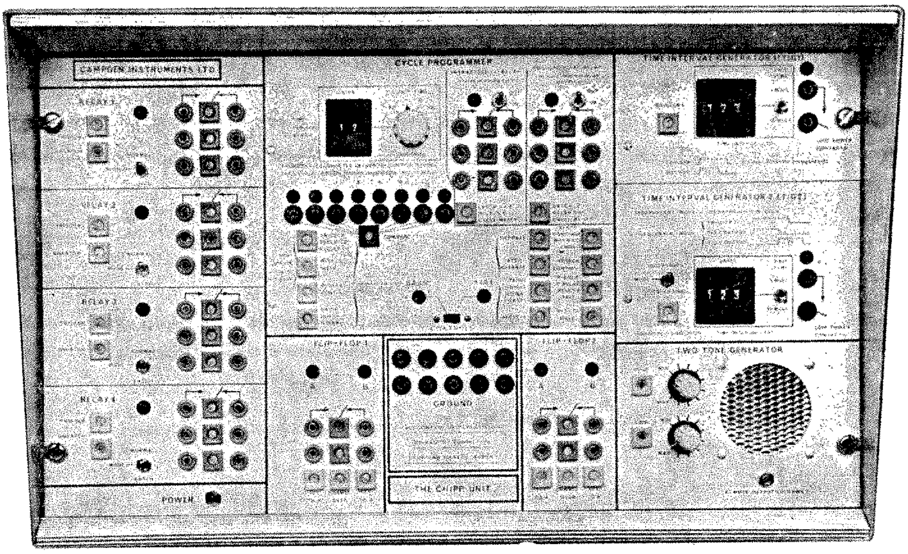
Wells, L. E. & Marwell, G. *Self-Esteem. Its Conceptualization and Measurement.* London: Sage. 1976. Pp. 290. £5.50; paper, £3.00.

Willis, J. W., Crowder, J. & Willis, J. *Guiding the Psychological and Educational Growth of Children.* Springfield, Ill.: Thomas. 1975. Pp. xi + 215. $11.50; paper, $6.95.

Wilmot, W. W. *Dyadic Communication: A Transactional Perspective.* London: Addison-Wesley: 1975. Pp. xv + 196. £3.00.

Wirt, R. D., Winnokur, G. & Roff, M. (eds). *Life History Research in Psychopathology.* Vol. 4. Minneapolis: University of Minnesota Press. 1976. Pp. ix + 360. $19.50.

NEW from CAMPDEN INSTRUMENTS – a powerful and unique package of control logic – THE CHIPP UNIT (Controller for Human Information Processing Psychology)



The CHIPP UNIT has been designed by CAMPDEN in consultation with Dr. John Frisby and Dr. Adrian Simpson of the Department of Psychology in the University of Sheffield who have also written the comprehensive Teaching Manual. The design of the CHIPP UNIT caters ideally for the present day trends in Psychological Experimentation thus superseding existing "Student Laboratory Units"

SPECIAL FEATURES INCLUDE:–

- LOGIC FACILITIES IN A SINGLE INSTRUMENT
- COMPLETE ELECTRICAL ISOLATION BETWEEN THE CHIPP UNIT AND OTHER EQUIPMENT
- L.E.D. INDICATOR LAMPS TO SHOW STATES AT ANY GIVEN INSTANT
- SIMPLE LOGICAL OPERATION BY GROUND SIGNALS
- CLEAR PANEL DESIGN FOR EASE OF OPERATION
- IDEALLY SUITED FOR STUDENT USE AS WELL AS RESEARCH – INCLUDING ANIMAL WORK

The CHIPP UNIT comes with a comprehensive USER'S MANUAL which includes a full description of how it is operated. This Manual includes illustrations on the programming of a number of psychological experiments and is written with the student user particularly in mind so that students should be able to teach themselves how to use it.

For more information on the CHIPP UNIT also on our FISH experimental units write to:–

CAMPDEN INSTRUMENTS LTD.,
186, CAMPDEN HILL ROAD,
LONDON, W8 7TH.

Telephones:– (01) 727 3437
(01) 229 3636

Oxford Books for Students

Principles and Methods of Psychology

Robert B. Lawson, Steven G. Goldstein and Richard E. Musty

This book provides an introduction to the basic areas of psychology which include learning, motivation, perception, social psychology, individual differences, and experimental design and statistics. It is intended for use in experimental courses and in introductory courses where the emphasis is on experimental psychology. A teacher's manual with multiple choice test items is available on request from the University Promotion Section. Illustrated £8.50

Experiments in Psychology

Robert B. Lawson, Steven G. Goldstein and Richard E. Musty

The purpose of this laboratory manual and research kit is to provide the student with practical experience in conducting and reporting psychological experiments in the areas of perception, learning, motivation, and social psychology and individual differences. The manual contains some background information and procedural instructions to carry out the experiments in each of the above areas while the laboratory kit contains the requisite apparatus to complete each experiment. Illustrated £5.95

The Experience of Anxiety

A Casebook

Edited by Michael J. Goldstein and James O. Palmer

The second edition of this collection of case histories in abnormal psychology includes nine new cases covering phobia, hyperkinesia, death, suicide, female identity conflict, ethnic identity conflict, anorexia nervosa, drug abuse, and divorce. The new cases put more emphasis on the problems of young people and also on family interaction. Second edition paper covers £3

A Handbook for the Study of Suicide

Edited by Seymour Perlin

This book views the phenomenon of suicide from a historical, literary, and philosophical background. It then moves by way of the social sciences to a balanced psychiatric onsideration of suicidal behaviour and presents the contribution that each relevant iscipline makes to the study of suicide. It also offers a modern perspective on rational otivation toward suicide which diverges from the traditional psychiatric position hat all suicides represent psychopathology. Illustrated £5.75 paper covers £3.45

nspection copies of these books are available to lecturers wishing to consider them for doption as required texts. Please write to University Promotion Section, Oxford niversity Press (BJP), Ely House, 37 Dover Street, London W1X 4AH, with details of he course and the number of students involved.

REVUE DE PSYCHOLOGIE APPLIQUEE

Publication Trimestrielle

Directeurs:

Dr P. PICHOT et P. RENNES

1er Trimestre 1976 – Volume 26 – No. 1

SOMMAIRE

C. LEVY-LEBOYER, G. MOSER, B. VEDRENNE et M. VEYSSIERE – Le diagnostic du style d'encadrement: la signification du LPC de Fiedler.

A. M. MEAR-CRINE – La modification du comportement en milieu scolaire.

A. ZIV – Le problème des enfants intellectuellement surdoués.

S. ROSENZWEIG – L'agression et le Test de Frustration de Rosenzweig.

J. WOLPE – Portrait, notice biographique et bibliographique.

Rédaction et Administration:

48, avenue Victor HUGO – 75783 PARIS CEDEX 16

Abonnements

(un an) France: 45 F; Étranger: 55 F

Le numéro France: 12 F; Étranger: 12 F

Règlements au nom des Éditions du Centre de Psychologie Appliquée

Numéro spécimen sur demande

WILEY

LONDON – NEW YORK – SYDNEY – TORONTO

ADVANCES IN PSYCHO-BIOLOGY, Volumes 1 & 2

edited by **Grant Newton and A. H. Riesen,**
both of *University of California, Riverside.*

The first volume deals with broad principles and implications in areas of growth, development, and longevity – all relevant to problems of behaviour. The second volume reviews current studies of early growth and development, plasticity of the nervous system, development of visual perception, and the relationship between the early environment and learning abilities.

316 pages (Vol. 1) May 1972 $27.00 £15.05
342 pages (Vol. 2) December 1974 $29.70 £16.55

BRAIN, BEHAVIOUR AND DRUGS

by **D. M. Warburton,**
Department of Psychology, Reading University.

Reflects a trend of psychopharmacology towards the explanation of behaviour in terms of the basic neurochemical systems in the brain.

290 pages December 1975 $14.25 £6.95

RECALL AND RECOGNITION

edited by **J. Brown,**
University of Bristol.

This is the first book in which the relation of recall to recognition is systematically examined and also the first in which recognition memory is given due prominence.

286 pages February 1976 $19.95 £9.75

THEORETICAL AND EXPERIMENTAL BASES OF THE BEHAVIOUR THERAPIES

edited by **P. Feldman** and **A. Broadhurst,**
Birmingham University.

The purpose of this volume is to identify and illustrate the links between experimental psychology and behaviour therapy and to look critically at some of the methodological and social issues in the practice of behaviour therapy.

474 pages April 1976 $24.55 £12.00

RESEARCH IN PSYCHOPHYSIOLOGY

edited by **P. H. Venables,**
University of York,
and **M. J. Christie,**
University of London.

A collection of original essays by international contributors which describes applications of psychophysiological methods to research areas which range from experimental psychology to psychiatry.

454 pages October 1975 $24.55 £12.00

John Wiley & Sons Ltd.
Baffins Lane, Chichester, PO19 1UD,
Sussex, England

European Monographs in Social Psychology Number 7

Series editor: H. Tajfel

Speech Style and Social Evaluation

H. Giles and P. F. Powesland

December 1975, xiv+218 pp., £6.60/$17.00 0.12.283750.9

This stimulating study presents a thorough investigation into the social importance of individual speech style. The authors examine the results of extensive research and show how listeners make judgements about a person's social status, competence and personality on the basis of his manner of speech.

"This work by Giles and Powesland comes as a breath of fresh air to the fields of social psychology and sociolinguistics. For me its major impact lies in the fact that it casts a new light on the social significance of the more superficial or surface aspects of human relationships ... Giles and Powesland show us how very important 'surface stuff' of this sort actually is ... The research throughout is done with humour and freshness and an unbelievable amount of ingenuity. G. B. Shaw would have enjoyed it — and learned from it."

Wallace E. Lambert
Professor of Psychology, McGill University, Montreal, Canada

Academic Press
London New York San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers
24–28 Oval Road, London NW1, England
111 Fifth Avenue, New York, NY 10003, USA
Australian Office
PO Box 300, North Ryde, NSW 2113, Australia

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA

rivista trimestrale pubblicata a cura dell'Università Cattolica del Sacro Cuore

direttore Leonardo Ancona

comitato di redazione

Mario Bertini – Renzo Carli – Fabrizio Napolitani – Luigi Pizzamiglio
Corrado Pontalti – Carlo Saraceni – Ciro Sdino

Anno XXXVI Ottobre–Dicembre 1975 Fasc. IV

Editoriale
Cornoldi, C. – Longoni, A. M., Memoria sequenziale, codifica verbale e per immagine
Mantovani, G., Fatica mentale industriale e studi sulla vigilanza
Ancona, L., La gruppoanalisi della omosessualità
Corradi Fiumara G., Psicoanalisi di una sindrome omosessuale
Corona, G. L. – Fenoglio, L. – Pinelli, P. – Zerbi, F., Sulla modalità d'azione dell'amitriptilina nei depressi: premesse concettuali e criteri di valutazione
Pinelli, P. – Tartara, A., L'elettroshockterapia unilaterale e bilaterale. Soglia per l'accesso bilaterale
Bonetti, A., Lo psicodramma breve. La drammatizzazione nelle psicoterapie brevi di gruppo

NOTE E DISCUSSIONI

Roth, M., Coping with the emotional sufferers
Siri, G., Verso la costruzione di una tipologia familiare

RECENSIONI

Ingrassia, G. e Coll., Fenomenologia omosessuale (R. Menarini)
'Letture da Le Scienze.' Problemi di sociologia e di demografia (L. Verdi Vighetti)
Nash, J., Psicologia dello sviluppo (G. Siri)
Porta, V., Neurologia illustrata (M. Cagossi)
Studies on excitation and inhibition in the retina. A collection of papers from the Laboratories of H. H. Hartline (L. Rositani Ronchi)
Sommario generale dell'annata

Redaction – Administration: 1, Largo A. Gemelli – 20123 MILAN (Italy).
Annual subscription: Italy L. 11.000 ; foreign countries L. 16.000 (or equivalent foreign money).

Environmental Psychology 2nd edition

By Harold M. Proshansky, William Ittelson and Leanne Rivlin, all of the City University of New York.

An extensive book of readings covering the significant issues relating to environmental psychology, and including theoretical, empirical and descriptive material. Its multidisciplinary approach and extensive bibliographies contribute to an up-to-date overview of the field.

£9.00 *640pp* *Hardback* *1976* *ISBN 0 03 089679 7*

Introduction to Personality 2nd edition

By Walter Mischel, Stanford University.

This updated edition of Walter Mischel's book provides a thorough discussing of the major theoretical approaches to personality assessment and therapy, plus an overview of current issues and findings in the study of personality.

£5.50 *560pp* *Hardback* *1976* *ISBN 0 03 089537 5*

The Enterprise of Living 2nd edition

By Robert W. White, Harvard University (Emeritus) Cambridge, Massachusetts.

This descriptive overview of the study of personality combines experimentation and theory, and draws upon case histories, clinical studies and other people's adaptive efforts, problems and life situations to increase the student's own self-understanding.

£8.50 *576pp* *Hardback* *1976* *ISBN 0 03 089522 7*

Abnormal Psychology: Changing Conceptions 2nd edition

By Melvin Zax PhD, and Emory L. Cowen PhD, Both of the University of Rochester, Rochester, New York.

Using an historical approach, this text stresses our evolving conceptions of abnormal psychology from traditional diagnosis and treatment to the latest developments in preventive, community-based psychology.

£10.00 *656pp* *Hardback* *1976* *ISBN 0 03 089517 0*

Fundamentals of Experimental Psychology 3rd edition

By Charles L. Sheridan PhD, University of Missouri, Kansas City, Missouri.

Using a simplified approach, this book presents a comprehensive range of basic methodological issues in experimental psychology, covering both human and animal examples.

£8.50 *512pp* *Hardback* *1976* *ISBN 0 03 083538 3*

Write now to Colin Dann, Publicity Manager, for your inspection copy, quoting reference HRW 17. Our Freepost address is Holt-Saunders Ltd., Freepost EA 86, Eastbourne, East Sussex. BN21 1BR.

Holt Rinehart & Winston

1 St. Anne's Road, Eastbourne, East Sussex BN21 3UN Tel: (0323) 638221

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
The Journal of Occupational Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, The British Journal of Medical Psychology, The Journal of Occupational Psychology and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1976 subscription prices, payable in advance are: £20.00 net per volume post free (US $60.00 in the USA and Canada) for *The British Journal of Psychology*; £16.00 net (US $48.00) for *The British Journal of Medical Psychology*; £13.00 net (US $39.00) for *The Journal of Occupational Psychology*; and £19.00 net (US $57.00) for *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian inquiries should be addressed to the Cambridge University Press, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15.00 per volume post free ($34.00 in the USA). Subscriptions should be sent to the Society's Agents, John Wright & Sons, Ltd., 42–44 Triangle West, Bristol BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £6.00 a year post free, payable in advance, or £2.00 each number. The subscription price in the USA is $20.00. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. D. B. Clarke, Department of Psychology, The University, Hull HU6 7RX;

for the *British Journal of Medical Psychology* should be sent to

Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London SW20 0NE;

for *The Journal of Occupational Psychology* should be sent to

Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster LA1 4YF;

for *The British Journal of Social and Clinical Psychology* should be sent to

Dr Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ: or to
Dr H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester M20 8LR;

for *The British Journal of Educational Psychology* should be sent to

Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable price 50p net (USA $2.00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18.00 (£16.00 if paid by Banker's Order). They are entitled to purchase any or all of the Society's Journals at special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London W1X 4DN.

The British Journal of Psychology *Volume 67, Part 3, August 1976*

CONTENTS

ZANGWILL, O. L. *page* 301
Thought and the brain

RICHARDSON, JOHN T. E.
The effects of stimulus attributes upon latency of word recognition 315

UNDERWOOD, GEOFFREY
Semantic interference from unattended printed words 327

SALTER, D., SPRINGER, G. AND BOLTON, L.
Semantic coding versus the stimulus suffix 339

LOHR, JEFFREY M.
Concurrent conditioning of evaluative meaning and imagery 353

THORNTON, E. W. AND HAGAN, PATRICIA J.
A failure to explain the effects of false heart-rate feedback on affect by induced changes in physiological response 359

CARROLL, D. AND SURTEES, P. G.
The electrodermal component of the orienting response in blind and deaf individuals 367

WOLFF, J. G.
Frequency, conceptual structure and pattern recognition 377

MOHAN, VIDHU AND KUMAR, DALIP
Qualitative analysis of the performance of introverts and extraverts on standard progressive matrices 391

FRASER, C. O.
Cognitive strategies and multidimensional scaling 399

MYERS, D. H.
Right- and left-handed counting of Braille dots in subjects unaccustomed to Braille 407

LIND, P. M.
The behavioural effects of single and combined stressors: a test of arousal theory 413

HARRIS, MARGARET
The influence of reversibility and truncation on the interpretation of the passive voice by young children 419

SYKES, DONALD H.
Stimulus processing and recognition memory in children 429

BOOK REVIEWS 439

OTHER PUBLICATIONS RECEIVED 457

© The British Psychological Society 1976

10 DEC 1976

Printed in Great Britain at the University Printing House, Cambridge

Volume 67, Part 4 November 1976

THE BRITISH JOURNAL OF
PSYCHOLOGY

Edited by A. D. B. CLARKE
with J. BROWN, R. CONRAD, N. F. DIXON,
G. W. GRANGER, I. MARTIN, R. L. REID,
B. TIZARD *and* P. B. WARR

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON NW1 2DB
32 EAST 57TH STREET NEW YORK, N.Y. 10022

Assistant Editor: D. I. WILLIAMS
Book Review Editor: D. C. KENDRICK
Press Editor: MICHELE BENJAMIN

Preparation of Papers for The British Journal of Psychology

In submitting papers for publication in *The British Journal of Psychology* contributors are asked to observe the following requirements.

1. Articles must be as short as adequate presentation allows. It is rarely possible to publish papers more than 7000 words in length, and most papers suitable for publication in the *Journal* gain by being much shorter than this.

It may be possible occasionally to give early publication to very short articles, not exceeding two printed pages in length (about 1000 words). As a rule such articles need not have a summary or titled headings.

2. All papers and reviews must be typed on one side of the paper only with double spacing including list of references and all the pages numbered. Footnotes should be avoided. Articles should where possible be subdivided with titled headings, following the pattern of articles published from Vol. 56 (1965). A brief summary, suitable for use as an abstract should precede the article.

Manuscripts should be submitted in duplicate

3. The Editor cannot consider for publication articles seriously deficient in presentation. It is essential that a typescript submitted should represent the final form in which the author wishes the paper to appear. Authors may be charged for any proof changes (other than printers errors) that may involve re-setting of lines or paragraphs. Insertion of new material involving re-paging, and changes in blocks made from diagrams or drawings, are so costly that authors are well advised to regard them as out of the question.

4. Tables, photographs and figures which are costly to print should be used only when they save considerable space or are essential for effective presentation: and in no case should both a figure and table be used to convey the same information. Each table or figure should be numbered and on a separate sheet with its approximate position indicated in the typescript.

Drawings should be in Indian ink on heavy unruled paper or on graph paper ruled in light blue. Lettering on graphs, etc., is best indicated in a tracing placed over the diagram. Diagrams are normally reduced in reproduction and allowance must be made for corresponding reduction in the thickness of lines and the spaces between them. The numbered legend of a figure should be typewritten on a separate sheet of paper, not lettered on the drawing.

The object of the caption for a table or the legend for a figure should be to make the table or the figure comprehensible without reference to the text.

Authors must ensure that formulae are clear, particularly when these contain subscripts.

5. References in the text should be indicated by dates in parentheses after the author's names and corresponding to an alphabetical list at the end of the paper. This list must give author's names with all the details and the layout illustrated in the following examples:

ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. abnorm. soc. Psychol.* **26**, 231–284.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*. London: Macmillan.

RÉVÉSZ, G. (1951). Le problème du génie. *Anée psychol.* **50**, 83–96.

World list abbreviations are used for journals.

The Oxford Dictionary is taken as the authority on spelling.

The British Psychological Society subscribes to the Fair Copying Declaration of the Royal Society under the terms of which reproduction of material published in this *Journal* is permitted.

© The British Psychological Society 1976

Permissions

For permission to reproduce material from *The British Journal of Psychology*, please apply to the London or New York office of Cambridge University Press.

ISI Tear Service, 325 Chestnut Street, Philadelphia, Pennsylvania 19106, USA is authorized to supply single copies of separate articles for private use only.

Subscriptions

The British Journal of Psychology is published quarterly.

The subscription price (including postage) of volume 67, 1976, is £20.00 net (US $60.00 in the USA and Canada). Single parts cost £6.00 each (US $18.00 in the USA and Canada).

Second class postage paid at New York, N.Y.

Published by Cambridge University Press, Bentley House, 200 Euston Road, London NW1 2DB and 32 East 57th Street, New York, N.Y. 10022.

Br. J. Psychol. (1976), 67, 4, *pp.* 461–466
Printed in Great Britain

A RELATIONSHIP BETWEEN EYSENCK'S P SCALE AND CHANGE IN MUSCLE ACTION POTENTIALS WITH ATTENTION TO PERCEPTUAL TASKS

By JOHN HINTON and BRIAN CRASKE
Department of Psychology, University of Southampton

An experiment is described in which an index of change in muscle tension produced by perceptual discrimination tasks, and scores on Eysenck's PEN scales, were measured on 39 university students. It was hypothesized that if P is an index of psychoticism, as has been claimed (Eysenck & Eysenck, 1968) then an index of one central feature of so called 'psychotics', namely attention difficulty while decision making, should correlate positively with P.

P was found to correlate significantly with a measure of increase of EMG on the perceptual discrimination tasks ($r = 0{\cdot}54$, $P < 0{\cdot}001$). Correlations of this index with N and E were zero. The results are considered to be particularly significant, since scores on P, E and N gave almost zero intercorrelations in the sample tested, and L scores were exceedingly low. Further evidence is provided which suggests that both P and muscle tension increase are directly related to the degree of endorsement of paranoid ideas.

Eysenck & Eysenck (1968*a*) have claimed that P is a discrete dimension of personality in the normal population. However, there is some evidence that P is not entirely independent of N (Eysenck & Eysenck, 1968*b**; Davis, 1974), and in a paper by Verma & Eysenck (1973) the correlation between these scales is particularly high. Since many of the items in the P scale relate to paranoid and hostile attitudes, and lack of social conformity, they would not appear to be obviously associated with neuroticism. Eysenck maintains that P is a unitary dimension of psychoticism, but this view is open to criticism on the grounds that the label 'psychotic' covers a number of different behavioural abnormalities (Venables & Wing, 1962; Venables, 1957; Inglis, 1966). However, there is considerable research evidence that abnormalities of attention, perception and thinking, are primary features of patients labelled 'psychotic' (Payne, 1964; Payne, Caird & Laverty, 1964; McGhie & Chapman, 1961; Venables, 1963, 1964). Therefore, if P is a measure of psychoticism, as proposed by Eysenck, it should be directly related to a measure of difficulty in paying attention on simple perceptual tasks.

Research by Eason & White (1961) indicates that what may be called 'attentional effort' is positively correlated with the magnitude of action potentials in those muscles which are not involved in the tasks being undertaken. They provide evidence for a general EMG factor – each subject having his own unique 'muscular tension personality'. This was previously suggested by Malmo, Shagass & Davis (1951*a*). In an experiment on attentive listening, Wallerstein (1954) obtained results which indicate that magnitude of increase in the EMG directly relates to the degree of interest in prose passages: first presentations resulted in greatest muscle tension. The 'attentional effort' hypothesis is also consistent with the findings of Davis (1948) and Benson & Gedye (1961).

The evidence, therefore, suggests that changes in EMG reflect degree of effort to attend, this being dependent upon the perceived difficulty of the task and degree of motivation to perform the task. The research of Eason & White (1961) shows that recording from the neck trapezius muscle gives the best single index of generalized motor action potentials in 'irrelevant' muscles.

---

* Eysenck has more recently introduced the P.Q. which is claimed to give no significant correlation between P, E and N.

Very little research has been published on the relationships between EMG and groups of individuals labelled 'psychotic'. Goldstein (1965) showed that patients diagnosed as psychotic or depressive have abnormally high EMG reactivity relative to normals and other groups of patients. Malmo *et al.* (1951*a*,*b*) noted that relatively larger increases in muscle tension occurred in patients classified as chronic withdrawn psychotics, when they were given tasks involving discrimination and decision making, as compared with passive stress tests. Pain–stress tests produced maximum increases in EMG in patients labelled acute psychotic and psychoneurotic, compared with a normal control group. The psychoneurotics included depressives and patients reporting pathological tension.

Here it is hypothesized that, if Eysenck's P scale measures an aspect of psychotic behaviour relating to attention difficulty, then P score should be positively correlated with an EMG measure of attentional effort on perceptual discrimination tasks.

METHOD

Subject selection

Thirty-nine naive subjects, 12 of whom were females, were obtained from a university student population. Ages ranged from 19 to 30 years. Nineteen of these were selected at random from students taking first-year psychology courses, and there were no refusals. The rest were drawn randomly from the University population, with 70 per cent refusals. All subjects were paid, and they were informed that payment would be made at the end of the third and final test session. They were selected using the following criteria: age less than 30, no known defects of vision or audition, of satisfactory general health. No drugs or alcohol were to be taken on the day of test.

General laboratory setting

A quiet test booth was used. The subject was placed in a comfortable semi-reclining position with legs supported. The temperature was controlled at 25 °C (± 1 °C) and the booth was dimly lit by a deep red light situated behind the subject. The experimenter controlled the test sequence from an adjacent room. The experimenter could speak to the subject via earphones and the subject could signal to him by pressing a button. This simultaneously activated a buzzer and operated a marker on the polygraph recording chart.

EMG recording

A Devices 2-Channel Recorder was used. Motor action potentials were amplified with the AC7 pre-amplifier set to maximum gain and time constant of 0·3 sec. A DC5 power amplifier was used. No filtering of high frequencies was employed. EMG was integrated using a two second epoch.

Electrode placement

After shaving and cleaning the neck, a pair of 10 mm silver disc electrodes were attached to the skin over the right neck trapezius muscle. Pick up of skin potential responses and ECG artifacts were reduced by the use of concentric earth electrodes round each pick-up electrode. Mains hum artifact was eliminated by having an earthed wire mesh cage around the subject, double screening all leads, and reducing of electrode–skin contact resistance to below 6 kilohm.

The perceptual tests

A series of decision-making tasks was presented to the subject, involving auditory and visual discrimination, and utilizing the phenomenon of intrasensory inhibition (von Bekesy, 1967). These were given in sets of four trials for each type of test. Each batch of tests was presented four times, using a balanced alternating type of design to avoid fatigue (cf. Hinton, 1975). The tests were as follows:

(1) *Auditory.* (*a*) Low-frequency noise pulses (40–300 Hz) were delivered over earphones, with interaural time delay being varied by experimenter. A decision was required on lateralization of the sound, i.e. when the sound just reached the side of the head, as it appeared to move laterally. (*b*) A similar judgement was required using high-frequency noise (8–14 kHz), but in this case interaural intensity difference was the independent variable.

(2) *Visual.* A brightness matching task was used. The subject was required to judge when the brightness of a disc in the central area of a display matched in brightness an annular surround. Experimenter controlled the intensity of the latter.

Tests were carried out on three consecutive days, between 14.00 and 16.00 hr. Before and after the test session a period of relaxation was allowed. It was assumed that on the first testing session subjects would not be fully relaxed, and test procedures were being learned. Thus measurements of EMG were analysed for the second and third testing sessions.

Personality test administration

The EPI (form A) was administered at the end of the second session and the PEN scale was given after the third one. Subjects were told that, for the tests to give valid and useful results, it was essential to complete the questionnaires *very* quickly: they were advised to endorse the first answer which came to mind, aiming to complete the test within 3 min. Care was taken to present the tests as entirely confidential.

Collection of data

The integrated EMG was recorded during the 4 sec prior to making decisions on all discrimination tests, and the mean was calculated. The overall mean of integrated EMG was taken – combining results from all tests, since no differences were observed in responses to the various audio and visual tasks. The minimum levels of integrated EMG were also recorded. For the correlational analysis, measures were taken from session 2 and 3. It was considered that, by expressing EMG on test as a function of the minimum tonic level, inter-subject differences in the number of effective motor units would be eliminated, i.e. the electrode placement error would be minimized. Test–retest reliabilities were conducted, and only reliable EMG measures analysed. Relationships between PEN scale scores and change in EMG on paying attention, were investigated by correlational analyses. Also, high P and low P scoring groups were compared with their respective EMG changes over consecutive sessions.

RESULTS

Reliability of integrated EMG measure

Comparing session 2 with session 3, the reliability of the measure of increase in muscle action potentials, expressed as a percentage of minimum tonic level, yields $r = 0{\cdot}66$ ($P < 0{\cdot}001$ for n of 39). This degree of correlation is satisfactorily high for use in further analyses.

The means and standard deviations of the PEN scale scores in the student sample are similar to those reported by Eysenck & Eysenck (1968*b*) for E and N. However, scores on P are exceptionally high in this experimental group (mean 4·03, S.D. 1·94), while L scores are very low (mean 2·00, S.D. 1·30).

An observation of principal importance is that the three PEN scale scores were almost uncorrelated (Table 1). Thus meaningful comparisons can be made between personality types and relationships between the psychophysiological variable and personality dimensions can be more clearly interpreted. Correlations between PEN scale scores and muscle tension change are shown in Table 1. Percentage increase in muscle action potentials, on attending to simple perceptual tasks, is significantly correlated with P score. There is zero correlation of the EMG index with E and N.

Table 1. *Correlation matrix of PEN scale scores and EMG percentage change with attention*

	EMG percentage change	P	E	N
EMG percentage change				
P	+0·537			
E	−0·047	+0·057		
N	+0·087	−0·122	−0·051	

$n = 39$; for $P < 0{\cdot}001$, $r > 0{\cdot}50$.

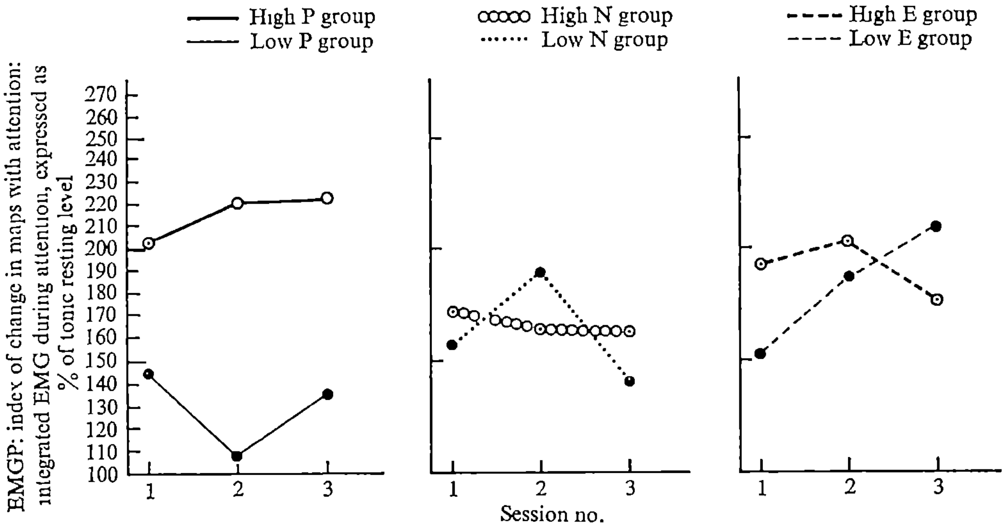


FIG. 1. Comparison between personality types on muscle action potential changes with attention, illustrating effects of repeating test sessions on separate days. (High and low P group mean scores for sessions 1, 2 and 3 are significantly different: $P < 0{\cdot}02$; analysis of variance.)

To allow the examination of EMG percentage change scores for different personality groups, over repeated test sessions, the subjects were separated into high and low scorers on each of the PEN scales, thus:

High P group: P score $\geqslant$ 1 s.d. above mean $n = 13$.
Low P group: P score $\leqslant$ 1 s.d. below mean $n = 10$.

Extraverts: E score $\geqslant$ 1 s.d. above mean $n = 8$.
Introverts: E score $\leqslant$ 1 s.d. below mean $n = 9$.

High N scores: N score $\geqslant$ 0·8 s.d. above mean $n = 9$.
Low N scores: N score $\geqslant$ 0·8 s.d. below mean $n = 8$.

The graphs in Fig. 1 give the required comparisons between groups, showing change over consecutive days.

As would be expected from the correlational analysis, high P scorers show a significantly greater increase in muscle tension on paying attention, compared to low P scorers. However, this EMG index shows progressive increase on repeated testing with the high P group. High and low scorers on N and E scales are not differentiated by the EMG index. Introverts show rising percentage EMG increase over three sessions on consecutive days, but other groups show no clear trends.

A test was carried out to check whether the results held for both male and female subjects. Product-moment correlations between the EMG index and P score were as follows:

Males $(n = 27)$, $r = +0{\cdot}560$ $(P < 0{\cdot}01)$
Females $(n = 12)$, $r = +0{\cdot}436$ (n.s.)

In interpreting these results it should be noted that no females scored above five on the P scale.

DISCUSSION

The results are consistent with the view that P is a dimension of personality which is distinct from E and N. As hypothesized, P score related directly to increase in generalized muscle action potentials on attending to simple perceptual discrimination tasks.

One point, not given *a priori* consideration, is that no feedback was possible on the discrimination tasks. This fact could affect interpretation of the results. Attention difficulty in high P scorers may have resulted from a distracting paranoid 'threat' reaction to a situation where subject does not know whether his responses are right or wrong. Ideally a follow-up test is required, including a task with feedback. However, it was decided to check whether high P scores resulted from endorsement of obviously paranoid P items. It was found that P score correlated 0·68 $(P < 0{\cdot}001)$ with the 'scale' of face valid paranoid items. This result is not so surprising when one considers the report of Verma & Eysenck (1973) which showed clearly that P correlated very significantly with measures of paranoia and hostility.

Finally, it was found that the paranoid item 'scale' correlated 0·49 $(P < 0{\cdot}01)$ with the EMG index. Since lie scale scores were extremely low (< 4 on the student sample), it is reasonable to assume that those with paranoid response tendencies were not defensive in answering paranoid items on the P scale. Thus, individuals with paranoid attitudes probably show greater increases in muscle tension. Further validation of the above points is provided by observations of high P scorers during test sessions: the records showed that the behaviour of these individuals tended to be paranoid and hostile. These indications are also compatible with the speculation of Gray (1971) that P is a dimension of personality relating to activity level in the amygdaloid flight–fight system, which he claims is a sex-linked characteristic. The results of this study support his view, since female P scores and EMG change scores were relatively low. This could explain the lack of a statistically significant correlation between P and the EMG index in the female sample.

It may be concluded that Eysenck's P scale measures paranoid reaction tendencies which are linked to attention difficulty – this being reflected in increase of muscle tension on tests where decisions are required without feedback. Currently this conclusion is being tested by carrying out a repeat study on patients detained in a maximum security hospital.

REFERENCES

BEKESY, G. VON (1967). *Sensory Inhibition*. Princeton, N.J.: Princeton University Press.

BENSON, A. J. & GEDYE, J. L. (1961). Some supraspinal factors influencing generalized muscle activity. In Franklyn, Ward & Wheeler (eds), *Symposium on Skeletal Muscle Spasm*, pp. 31–50. Leicester.

DAVIS, H. (1974). What does the P scale measure? *Br. J. Psychiat.* **125**, 161–168.

DAVIS, R. C. (1948). Responses to 'meaningful' and 'meaningless' sounds. *J. exp. Psychol.* **38**, 744–756.

EASON, R. G. & WHITE, C. T. (1961). Muscular tension, effort and tracking difficulty: studies of parameters which affect tension level and performance efficiency. *Percept. mot. skills* **12**, 331–372.

EYSENCK, H. J. & EYSENCK, S. B. G. (1968*a*). A factorial study of psychoticism as a dimension of personality. *Multiv. Behav. Res.* (All Clinical Special Issue).

EYSENCK, S. B. G. & EYSENCK, H. J. (1968*b*). The measurement of psychoticism: A study of factor stability and reliability. *Br. J. soc. clin. Psychol.* **7**, 286–294.

GOLDSTEIN, I. B. (1965). The relationship of muscle tension and autonomic activity to psychiatric disorders. *Psychosom. Med.* **27**, 39–52.

GRAY, J. A. (1971). Causal theories of personality and how to test them. Paper read at 3rd Banff Conference on Theoretical Psychology.

HINTON, J. (1975). PhD thesis, University of Southampton.

INGLIS, J. (1966). *The Scientific Study of Abnormal Behaviour*. Chicago, Ill.: Aldine.

MALMO, R., SHAGASS, C. & DAVIS, J. L. (1951*a*). Electromyographic studies of muscular tension in psychiatric patients under stress. *J. clin. exp. Psychopath.* **12**, 45–66.

MALMO, R., SHAGASS, C. & SMITH, A. A. (1951*b*). Responsiveness in chronic schizophrenia. *J. Personality* **19**, 4.

MCGHIE, A. & CHAPMAN, J. (1961). Disorders of attention and perception in early schizophrenia. *Br. J. med. Psychol.* **34**, 103–116.

PAYNE, R. W. (1964). The measurement and significance of overinclusive thinking and retardation in schizophrenic patients. Paper presented to the American Psycho-pathological Association.

PAYNE, R. W., CAIRD, W. K. & LAVERTY, S. G. (1964). Overinclusive thinking and delusions in schizophrenic patients. *J. abnorm. soc. Psychol.* **68**, 562–566.

VENABLES, P. H. (1957). A short scale for rating activity - withdrawal in schizophrenics. *J. ment. Sci.* **103**, 197–199.

VENABLES, P. H. (1963). Selectivity of attention, withdrawal and cortical activation. *Archs gen. Psychiat.* **9**, 74–78.

VENABLES, P. H. (1964). Input dysfunction in schizophrenia. *Progr. exp. Person. Res.* **1**, 1–47.

VENABLES, P. H. & WING, J. K. (1962). Level of arousal and the subclassification of schizophrenia. *Archs gen. Psychiat.* **7**, 114–119.

WALLERSTEIN, H. (1954). An electromyographic study of attentive listening. *Can. J. Psychol.* **8**, 228–238.

VERMA, R. M. & EYSENCK, H. J. (1973). Severity and type of psychotic illness as a function of personality. *Br. J. Psychiat.* **122**, 573–585.

(*Manuscript received* 9 *May* 1975; *revised manuscript received* 18 *July* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 467–474
Printed in Great Britain

EXTRAVERSION, NEUROTICISM AND STRENGTH OF THE NERVOUS SYSTEM

BY JEAN-YVES FRIGON
Department of Psychology, University of Montreal, Canada

The hypothesized identity of the dimensions of extraversion–introversion and strength of the nervous system was tested on four groups of nine subjects (neurotic extraverts, stable extraverts, neurotic introverts, stable introverts). Strength of the subjects' nervous system was estimated using the electroencephalographic (EEG) variant of extinction with reinforcement. Introverted subjects were found to have weak nervous systems, according to the EEG index, while extraverted subjects had strong nervous systems, thus confirming the hypothesis. It was also found that the dimension of strength of the nervous system was unrelated to differences in neuroticism. The results are interpreted as adding support to Eysenck's theory relating differences in extraversion–introversion to differences in cortical arousal.

Since the publication of Gray's book (1964) on the dimension known in Russia as 'strength of the nervous system (NS)' and his reinterpretation of it as a dimension of arousability, few studies have tried to relate that dimension of individual differences to measures of personality. According to Gray (1967) and Eysenck (1966, 1967), the dimension of personality that would best correspond to the dimension of strength of the NS is that of extraversion–introversion since they both seem to be dependent on differences in arousability. Gray (1967), in an excellent analysis, has argued that the most plausible hypothesis is that 'the dimensions of strength of the nervous system and introversion–extraversion are identical, with the weak nervous system corresponding to the introvert' (p. 153), an hypothesis which is also adopted by Eysenck (1967).

Despite the fact that the hypothesis relating introversion–extraversion to strength of the NS is theoretically well founded, studies which have tried to validate it experimentally have produced discouragingly deceptive and contradictory results (Mangan, 1967; Mangan & Farmer, 1967; Zhorov & Yermolayeva-Tomina, 1972). Considering that these results all failed to confirm the correspondence between introversion–extraversion and strength of the NS, Zhorov & Yermolayeva-Tomina (1972) conclude 'that either (1) the hypothesis of the relation between introversion and weakness is, despite its plausibility, wrong, or (2) the method of determining strength employed in these experiments was not adequate for the task of measuring the strength of the nerve cells of the brain structures...' (p. 267) and they add that 'the latter assumption is probably closer to the truth'.

It is also our opinion that this latter assumption reflects the truth. The relationship between extraversion–introversion and strength of the NS has been studied with an estimation of nervous system strength obtained by the reaction time (RT) method (Mangan & Farmer, 1967; Zhorov & Yermolayeva-Tomina, 1972). This method, however, does not seem appropriate to such a study for two main reasons.

First, the RT method provides only an indirect index of nervous system strength. According to Nebylitsyn (1972, p. 165), there are now two well-validated methods which yield a direct index of the strength of the NS. These are the 'induction' and the 'extinction with reinforcement' methods, and the most feasible of the two is the electroencephalographic (EEG) variant of extinction with reinforcement (Teplov & Nebylitsyn, 1969; Nebylitsyn, 1972, pp. 157–168). It should be noted that Soviet researchers, when exploring the relationship between nervous systems differing in strength and their characteristic reactions to varying conditions, such as absolute sensitivity or sensory stimulation, have always used one or more direct referent indices of nervous system strength. However, those studies relating strength of the NS to introversion–extraversion (Mangan & Farmer, 1967; Zhorov & Yermolayeva-Tomina, 1972) have not employed a direct referent index of nervous system strength. This is unfortunate since the indirect index provided by the RT method may not be sensitive enough to differences in strength.

Second, although the index of nervous system strength obtained with the RT method is said to correlate positively with other indices of strength, this correlation has not been very well validated. Nebylitsyn (1972, pp. 191–208) has argued that the RT method can be used to estimate the strength of the NS, provided that one uses the ratios of latency for a given intensity to latency to the highest stimulus intensity, rather than the absolute magnitudes of response latency. This conclusion is based on two studies in which the RTs of two groups of subjects (strong and weak NSs) are measured under six levels of stimulus intensity and two conditions – non-caffeine and caffeine – (Nebylitsyn, 1960: see Nebylitsyn, 1972, p. 192). Using experimental design terminology, these situations define a $P \times Q \times R$ split-plot factorial design with repeated measures on the last two factors (Kirk, 1968, p. 298). However, it is evident, from the analysis of variance table presented by Nebylitsyn (1972, p. 200), that a $P \times Q \times R$ factorial design for independent measures on all three factors was used (Kirk, 1968, p. 217), and this analysis is inappropriate for these data.

The sums of squares of the main and interaction effects would be identical under both techniques of analysis. But, with the analysis of variance appropriate to this type of repeated measures design, the error term would be divided into four parts, the partitioning of the degrees of freedom would be different, and the F ratios would accordingly be computed differently. Without the raw data it is difficult to evaluate exactly how the use of the appropriate analysis design would have modified Nebylitsyn's results. Nevertheless, it is very likely that the between-subjects F ratio has been inflated due to the improper analysis. Consequently, the significant strength effect reported by Nebylitsyn may not really be significant. This inappropriate analysis may not invalidate completely the RT method. But it surely reveals an inadequate validation procedure and puts some doubts on its appropriateness as a measure of nervous system strength. Therefore, the use of this method may account for the contradictory results which have been obtained in studies on the relationship between extraversion–introversion and strength of the NS.

The term 'extinction with reinforcement' refers to a temporary decrement of a conditioned response (CR) which occurs through multiple presentations of the conditioned stimulus–unconditioned stimulus (CS–UCS) in rapid succession, a procedure which theoretically has an exhausting effect on the nerve cells. Briefly, the method consists in presenting subjects with a number of CS–UCS combinations in which the intertrial interval is shorter than under normal conditions. A decrease in the magnitude of the CR at the end of this series of rapid presentations of CS–UCS is taken as an indication of weakness of the NS; the absence of such a decrease or an increase in the magnitude of the CR is an indication of strength of the NS (cf. Gray, 1964; Nebylitsyn, 1972; and Teplov & Nebylitsyn, 1969, for a thorough discussion of the theoretical basis and a more detailed account of the method of extinction with reinforcement). Two types of CRs have usually been used with this method: the first is what the Soviet psychologists call the photochemical reflex (photochemical variant of extinction with reinforcement); the second is based on the conditioned EEG desynchronization of alpha waves (EEG variant of extinction with reinforcement). The latter variant is less time consuming and is preferred by Teplov & Nebylitsyn (1969) and Nebylitsyn (1972).

The purpose of the present study was to test the hypothesis of the correspondence between introversion–extraversion and strength of the NS. The referent index of nervous strength was the EEG variant of extinction with reinforcement. It was hypothesized that introverts would show a decrease in the magnitude of the CR during extinction with reinforcement, and as such have a weak nervous system, while extraverts would show a stability or an increase in the magnitude of the CR, and thus have a strong nervous system. In statistical terms, we predicted an interaction between extraversion–introversion and measures of the magnitude of the CR during extinction with reinforcement.

METHOD

Subjects

The subjects were 36 male volunteers who were divided into four equal groups. They were all undergraduate students and were paid $3.00 for their participation. They were selected from a pool of 145 possible subjects on the basis of their scores on the Eysenck Personality Inventory (EPI), form A. The EPI was administered to all of the 145 students. Total mean scores and standard deviations were 11·75 and 3·10 respectively on the Extraversion (E) scale, and 9·68 and 3·85 respectively on the Neuroticism (N) scale. Experimental subjects were selected so that their scores on the test differed by more than one standard deviation from the mean on each of the four possible configurations of personality: high and low Extraversion (scores of 15 or more or 9 or less on the E scale) and high and low Neuroticism (scores of 14 or more or 6 or less on the N scale). There was no significant difference in N score between the two groups who differed significantly on E score (t = 0·348; d.f. = 34, n.s.); similarly, there was no significant difference in E score for the subjects high and low on N score (t = 0·579; d.f. = 34, n.s.).

Apparatus

Electroencephalographic activity was recorded on a Beckman–Offner Type R dynograph with AC/DC couplers. Two derivations were taken from the occipital region with Beckman–Offner silver/silver chloride skin electrodes (8 mm in diameter) filled with Beckman–Offner electrode paste.

During the experiment, subjects were seated in a dark soundproof room. Tones were presented through a speaker by a Heathkit sound generator. The intensity of the tones was measured by a Bruël and Kjaer sound level meter (meter model 2203; octave filter type 1613; microphone 4131).

Visual stimuli were presented from the exterior on a white screen superimposed on a one-way mirror that was fitted in one wall of the soundproof room. Grason–Stadler electronic timers and programming apparatus were used to control the duration of the CS, UCS, and interstimulus interval (ISI). Intertrial intervals (ITIs) were programmed on a film strip which activated the timers controlling the presentation of the stimuli.

The experimenter could communicate with subjects at all times through a two-way intercom.

The experimenter also monitored the subject and the stimuli from the outside, throughout the experiment, with a videoscope camera that was located inside the soundproof room.

Design and procedure

Upon entering the laboratory, the subject was seated in a comfortable armchair inside the soundproof room. General information about the apparatus and instructions about the nature of the experiment were given during the application of the electrodes. No information was given about the relationship between the CS and UCS. The subjects were tested in random order. The experimenter did not know, during testing, to which personality group they belonged.

Eleven of the subjects (six extraverts and five introverts) were tested in the morning and the remaining 25 subjects (12 extraverts and 13 introverts) in the afternoon. The effects of time of day on performance differences between introverts and extraverts are well known (Colquhoun & Corcoran, 1964). Since, however, an approximately equal number of extraverts and introverts were respectively tested in the morning and in the afternoon, any systematic effect of time of day on the conditioning data should have been prevented. Moreover, although time of day effects have been established with regard to subjects' performance, it is dubious whether alpha conditioning and indices of nervous system strength would be affected by this variable.

As usual in alpha conditioning studies, some subjects (four introverts and two extraverts) had to be rejected for various reasons, such as for not showing sufficient alpha waves or for being under heavy medication. These subjects were replaced with subjects from the original pool. They all met the criteria for subjects' selection set forth above.

The conditioning parameters were the following: the CS was a tone, 1000 Hz, 70 dB, of 5 sec duration. In most alpha-blocking conditioning studies, the UCS is usually a flash of light. The CR thus obtained, however, is very labile and the disappearance of the CR may be due to a decrease in novelty of the UCS, instead of being the result of extinction with reinforcement. To solve this problem, Teplov & Nebylitsyn (1969) and Nebylitsyn (1972) presented subjects with reinforcing stimuli which they called 'activating' stimuli. These stimuli were not merely lights, but slides of pictures with instructions to remember the details shown and to be prepared to answer questions in a post-experimental inquiry. Nebylitsyn reported that, with this type of reinforcement, alpha-blocking CRs were found to be very stable for most subjects. One can either use the same or different slides throughout the trials with comparable results (Nebylitsyn, 1972, p. 165). In the present study, such an 'activating' UCS was used and the same slide was presented in all trials. This slide was a representation of a pyramid in the Teotihuacan Valley (Mexico). The interval between CS onset and slide onset (ISI) was 1·5 sec. The slide was presented for 3·5 sec and both the CS and UCS ended simultaneously.

The conditioning and extinction with reinforcement paradigm called for the following conditions.

Habituation. First, the orienting reaction to the CS was extinguished with the tone presented alone until there was no alpha-blocking on three consecutive trials. The ITIs varied from 15–20 sec. A random number table was used to generate the sequence of ITIs. Sequences of ITIs also ranging from 15–20 sec were used during conditioning, differential conditioning, and CR baseline conditions, as well.

Conditioning. The conditioning paradigm consisted of 50 trials. On 45 occasions, CS–UCS were paired with an ISI of 1·5 sec. To observe the development of the CR, five CS-alone test trials were given on trials 10, 14, 22, 28 and 45. All 50 trials were given, even though the criterion of conditioning was an anticipatory response to the CS on three consecutive trials.

Differential conditioning. After the 50 conditioning trials, which in all cases were more than necessary to produce a stable CR, differential conditioning was begun, with a 500 Hz, 50 dB tone. Reinforced (CS–UCS) and unreinforced (differential stimulus) trials were presented according to the Gellerman table (Stevens, 1951, p. 533) and were continued until the differential stimulus did not produce alpha-blocking on three consecutive trials.

CR baseline. Before proceeding with the extinction with reinforcement series, 15 trials were given, 12 of these being reinforced trials (CS–UCS), and the other three being CS-alone test trials (trials 3, 11, 15). The mean duration of alpha-blocking in seconds on those three test trials was taken as the CR baseline prior to extinction with reinforcement. This measure of CR baseline is the only departure from Nebylitsyn's procedure, which measured it on the first three test trials of the extinction with reinforcement series. Since the ITI during that series of trials

Table 1. *Mean number of trials to reach the criterion of habituation, of conditioning and of differential conditioning*

	Habituation		Conditioning		Differential conditioning	
Group	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
E N	11·67	3·81	8·44	2·01	9·11	2·98
E S	15·22	6·94	8·22	2·22	10·00	5·36
I N	15·89	2·47	8·00	3·08	8·67	3·08
I S	13·56	4·61	8·56	2·01	8·44	2·88

is different from the ITI used during conditioning, it seemed preferable to have a measure of CR baseline with conditions identical to those prevailing during conditioning.

Extinction with reinforcement. In the extinction with reinforcement trials, subjects were presented with 48 CS–UCS pairings at ITIs of 4–5 sec, compared with 15–20 sec ITIs during conditioning. To observe the evolution of the CR, 12 test trials (CS-alone) were given. These occurred after trials 3, 10, 12, 19, 21, 24, 31, 34, 36, 39, 41 and 48. It should be noted that the ratio of reinforced/unreinforced trials by blocks of three successive test trials was the same here as it was during the measure of CR baseline. The ordering of the test trials also prevented the development of a temporal anticipation of CS-alone.

CR magnitude was the mean duration of alpha-blocking in seconds during the action of CS on blocks of three successive test trials (means of test trials 1–3, 4–6, 7–9, 10–12). Following Nebylitsyn's procedure, the mean of three successive test trials was taken to assure a more stable measure of CR magnitude, since there is always considerable trial-to-trial variability in EEG conditioning.

RESULTS

As reported by Teplov & Nebylitsyn (1969) and Nebylitsyn (1972), habituation, conditioning and differential conditioning, were fairly easy to obtain in all subjects. Means and standard deviations of the number of trials necessary to reach the criterion in each of these conditions are shown in Table 1. A 2 × 2 (high and low E and N) analysis of variance was performed on this variable, on each of the conditions of habituation, conditioning and differential conditioning. Neither of the F ratios computed in these three analyses revealed a significant effect of E or N. This indicates that the rates of habituation, of conditioning and of differential conditioning were approximately equivalent for all groups.

The main hypothesis of the present study predicted a significant E × blocks of trials interaction. This was tested with a 2 × 2 × 5 analysis of variance with repeated measures on the last factor. This analysis compared CRs magnitudes of the four personality groups (two levels of E and two levels of N) on the five successive blocks of three test trials (baseline was labelled as block 1). On the basis of the *a priori* hypothesis about the E × blocks interaction, a direct linear trend analysis of that interaction was performed (Kirk, 1968, p. 190). This analysis is presented in Table 2.

It can be seen that the linear trend of the E × blocks of trials interaction is significant. The CRs magnitudes over blocks of three test trials for E and I subjects are shown in Fig. 1. It appears that introverts are more cortically aroused than extraverts at the beginning of the blocks of test trials.

An analysis of the simple effects of blocks of test trials at each level of E reveals that introverts show a gradual decrease of the magnitude of the CR from blocks

Table 2. *Analysis of variance of CR magnitude (sec)*

Source	d.f.		m.s.	F
Between subjects	35			
E	1		26·07	7·22*
N	1		1·20	< 1
E × N	1		7·57	2·10
Subj w. groups	32		3·61	—
Within subjects	144			
Blocks	4		10·36	8·93**
E × blocks	4			
linear		1	6·32	4.42*
residual		3	0·43	< 1
Blocks × subj w. groups (linear)		32	1·43	—
N × blocks	4		2·54	2·19
E × N × blocks	4		0·55	< 1
Blocks × subj w. groups	128		1·16	—

$*P < 0{\cdot}05$; $**P < 0{\cdot}01$.

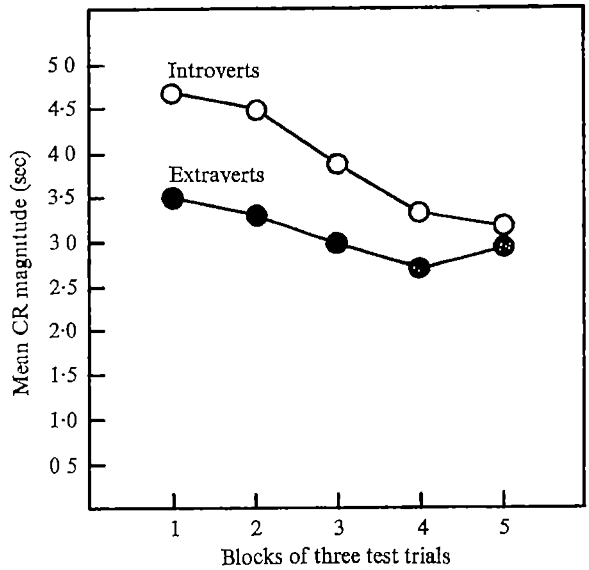


Fig. 1. Mean CR magnitudes over blocks of three test trials for extraverts and introverts.

1 to 5 ($F = 8{\cdot}47$, d.f. = 4, 128, $P < 0{\cdot}001$). However, the small decrease of the magnitude of the CR for extraverts is not significant ($F = 2{\cdot}10$, d.f. = 4, 128, $P > 0{\cdot}05$), thus explaining the significant linear trend of the E × blocks of trials interaction.

Table 2 also shows that Neuroticism did not affect CRs magnitudes either alone or in interaction with other factors.

DISCUSSION

The fact that extraverts and introverts did not differ in their rate of conditioning should not be taken as an argument against Eysenck's well-known view on the positive correlation between introversion and conditionability (Eysenck, 1965, 1967; Eysenck & Levey, 1972). Indeed, Eysenck (1967) and Eysenck & Levey (1972) have shown that, in conditioning studies, the parameters of stimulation must be carefully chosen, since it is possible to have conditions favouring introverted or extraverted subjects. The purpose of the present study was not to test the conditionability of the subjects, and it may be that the particular conditions of the experiment did not favour the better conditioning of one group or the other.

On the other hand, rates of habituation, of conditioning and of differential conditioning are related to another dimension of the neo-Pavlovian typology of individual differences, namely 'dynamism of the nervous system' (Gray, 1967; Teplov & Nebylitsyn, 1969; Nebylitsyn, 1972). Gray (1967) has already suggested that the dimensions of extraversion–introversion and dynamism of the NS are independent and the present results also lead us to this conclusion. It should be noted, however, that rates of habituation, of conditioning and of differential conditioning are only some of the many EEG indices of the dynamism of the NS (cf. Nebylitsyn, 1972, pp. 86–112). Therefore, it would be important to measure more of these indices in order to obtain a more valid estimate of the independence of these two dimensions.

The analysis of variance of CRs magnitudes showed that, in the 'extinction with reinforcement' condition, the duration of the conditioned EEG desynchronization progressively decreases for introverts and remains stable for extraverts. It is clear from the results that introverts are more cortically aroused than extraverts when both are exposed to objectively identical physical stimulation before proceeding with the series of rapid presentations of the CS. The EEG CR is unaffected in extraverts by the building up of excitation in the nerve cells during 'extinction with reinforcement', while introverts show a CR decrease. This is taken as an indication of weakness of the nervous system in introverts and strength of the nervous system in extraverts.

This finding suggests that Gray's (1967) and Eysenck's (1967) hypothesis on the identity between the dimensions of extraversion–introversion and strength of the NS is empirically sound. It should then be expected that extraverted subjects react, in general, as if they amplified the objective intensity of all stimulation and introverted subjects as if they damped it down (cf. Gray's analysis, 1964). Moreover, this indicates that most of the differences found between subjects differing in nervous system strength (cf. Gray, 1964; Nebylitsyn, 1972; Nebylitsyn & Gray, 1972) also apply to the extraversion–introversion dimension of personality. It has already been shown that introverts have lower auditory thresholds than extraverts (Smith, 1968). Shigehisa & Symons (1973) have also established that auditory sensitivity decreases in introverts, in reaction to an increase of the intensity of visual stimulation, and increases in extraverts. These results are consistent with data on the differences between individuals with strong and weak nervous systems. Gray (1964) suggests that the reactions of the weak vs. strong nervous systems individuals in relation to drugs, susceptibility to fatigue, etc., should also apply to the introversion–extraversion dimension of personality. These predictions are theoretically plausible, but they have yet to be empirically validated.

The results of this experiment may be interpreted as giving some support to Eysenck's theory linking differences in extraversion–introversion to differences in cortical arousal. Furthermore, these results also confirm Gray's (1967) view that the dimension identified in Russia as 'strength of the nervous system' is not related to Eysenck's measure of neuroticism.

As indicated earlier, the discrepancy between the present results and those obtained in previous studies could be attributed to the indirect nature of the index of nervous system strength, which probably does not allow a good discrimination among subjects on the dimension of strength of the NS. Thus, it can be seen that it is important to use a direct referent index of this dimension when exploring its relationship to other individual differences characteristics.

REFERENCES

COLQUHOUN, W. P. & CORCORAN, D. W. J. (1964). The effects of time of day and social isolation on the relationship between temperament and performance. *Br. J. soc. clin. Psychol.* **3**, 226–231.

EYSENCK, H. J. (1965). Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychol. Bull.* **63**, 258–270.

EYSENCK, H. J. (1966). Conditioning, introversion–extraversion and the strength of the nervous system. In V. D. Nebylitsyn (ed.), Symposium 9, *18th Int. Congr. Psychol., Moscow,* 33–44.

EYSENCK, H. J. (1967). *The Biological Basis of Personality.* Springfield, Ill.: Thomas.

EYSENCK, H. J. & LEVEY, A. (1972). Conditioning, introversion–extraversion and the strength of the nervous system. In V. D. Nebylitsyn & J. A. Gray (eds), *Biological Bases of Individual Behavior.* New York: Academic Press.

GRAY, J. A. (ed.) (1964). *Pavlov's Typology.* Oxford: Pergamon Press.

GRAY, J. A. (1967). Strength of the nervous system, introversion–extraversion, conditionability and arousal. *Behav. Res. Ther.* **5**, 151–169.

KIRK, R. E. (1968). *Experimental Design: Procedures for the Behavioral Sciences.* Belmont, Cal.: Brooks/Cole.

MANGAN, G. L. (1967). Studies of the relationship between Neo-Pavlovian properties of higher nervous activity and Western personality dimensions: IV. A factor analytic study of extraversion and flexibility, and the sensitivity and mobility of the nervous system. *J. exp. Res. Person.* **2**, 124–127.

MANGAN, G. L. & FARMER, R. G. (1967). Studies of the relationship between Neo-Pavlovian properties of higher nervous activity and Western personality dimensions: I. The relationship of nervous strength and sensitivity to extraversion. *J. exp. Res. Person.* **2**, 101–106.

NEBYLITSYN, V. D. (1972). *Fundamental Properties of the Human Nervous System.* New York: Plenum Press.

NEBYLITSYN, V. D. & GRAY, J. A. (eds) (1972). *Biological Bases of Individual Behavior.* New York: Academic Press.

SHIGEHISA, T. & SYMONS, J. R. (1973). Reliability of auditory responses under increasing intensity of visual stimulation in relation to personality. *Br. J. Psychol.* **64**, 375–381.

SMITH, S. L. (1968). Extraversion and sensory threshold. *Psychophysiology* **5**, 293–299.

STEVENS, S. S. (ed.) (1951). *Handbook of Experimental Psychology.* New York: Wiley.

TEPLOV, B. M. & NEBYLITSYN, V. D. (1969). Investigation of the properties of the nervous system as an approach to the study of individual psychological differences. In M. Cole & I. Maltzman (eds), *A Handbook of Contemporary Soviet Psychology.* New York: Basic Books.

ZHOROV, P. A. & YERMOLAYEVA-TOMINA, L. B. (1972). Concerning the relation between extraversion and strength of the nervous system. In V. D. Nebylitsyn & J. A. Gray (eds), *Biological Bases of Individual Behavior.* New York: Academic Press.

(*Manuscript received* 17 *June* 1975; *revised manuscript received* 12 *September* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 475–478
Printed in Great Britain

OCULOMOTOR ADAPTATION TO PRISMS: COMPLETE TRANSFER BETWEEN EYES

BY MARTIN CRAWSHAW* AND BRIAN CRASKE†
Department of Psychology, University of Southampton

It is argued that previous experiments investigating interocular transfer of adaptation have been concerned with an adaptation of the registered position of the limb used to indicate direction. Thus the question of interocular transfer of adaptation sited at the eye remains open.

Two experiments are reported. The first showed that there was complete transfer between exposed and unexposed eyes. The conditions were such that the adaptation could be explained only as a change in the registered position of the eye in the head, or of the head on the shoulders. The second experiment showed that changes of this latter kind did not occur.

Several experiments on interocular transfer of adaptation to laterally displacing wedge prisms have been reported. Generally, the finding has been that transfer is complete. However, these studies were performed relatively early in the current wave of interest in prisms, before it was accepted that there are at least two forms which this adaptation can take. These are: a change in the registered position of the eyes, and a similar change in registered position of the limb. Thus it happened that in every case the subject saw his arms in the exposure condition and also used them for pointing or reaching in the measurement condition. It is clear that if the registered position of the arm has undergone a change during exposure with one eye, the adaptation will be found when the same arm is tested in conjunction with the other eye. It is now clear that it was just this type of transfer that was found by Bossom & Hamilton (1963), Hamilton (1967), Pick, Hay & Pabst (1963) and Lund (1965). Similarly, if adaptation takes the form of a change in registered position of the head on the shoulders, a possibility suggested by Harris (1965), then a spurious interocular transfer will be found. In a theoretical article, Ebenholtz (1970) recognized the contaminant of using the arm for both exposure and measurement, and also the potential bias of measuring adaptation with the exposed eye before the occluded eye, while the after-effects are probably decaying.

Two other studies have dealt with the problem of transfer. They also used the arm in both exposure and measurement, but the exposures were lengthy (days rather than minutes) and it has been suggested (Hay & Pick, 1966) that lengthy exposure leads to adaptation of registered eye position. Kohler (1963, p. 35), pooling several measures of prism adaptation, reported that adaptation 'transferred to some degree' to the occluded eye. Hajos & Ritter (1965) reported that interocular transfer was complete; however, from their data, a comparison of post-treatment with pre-treatment conditions gives transfer of about two-thirds. There is thus some evidence from the only studies that might have involved adaptation of registered eye position that transfer between eyes was only partial.

In his article, Ebenholtz argued that, given the existence of adaptation of eye position, and given also Hering's law of equal innervation, there should be complete interocular transfer of this adaptation. While this seems plausible, it has not the status of logical necessity. It is known that each eye can adapt its registered position independently of the other. Thus using base-out prisms, Craske & Crawshaw (1974) showed quite clearly that the registered positions of the two eyes could be shifted in opposite directions. Hajos (1968) obtained similar results, but in this case it would be possible to hypothesize an adaptation of the limb contingent upon which eye was used. Foley & Miyanishi (1969) and Foley (1970), using the rather different paradigm of vertical displacement in opposite directions for the two eyes, found evidence for 'at least limited interocular independence'. Hence there is no direct evidence of the degree of transfer of adaptation of registered eye position to an occluded eye, and although Hering's law suggests it may be complete, evidence from related investigations indicates that this may not be so.

---

* Now at Department of Psychology, The University, Hull, HU6 7RX.

† Now at Department of Psychology, Memorial University of Newfoundland, St Johns, Newfoundland, Canada.

An experiment was performed to investigate the problem directly, without the contamination of limb adaptation. A second experiment was performed to investigate the existence of adaptation of the head on the shoulders.

It is important to note that the experiments avoided a hitherto unidentified source of bias arising from the use of a base-temporal prism before one eye and an occluder before the other. This could lead to the anomaly of apparent divergence of the two eyes if the exposed eye adapted but the occluded eye did not. However, with a base-nasal prism and no interocular transfer, the eyes will be judged to be in asymmetrical convergence which, although erroneous, is a perfectly normal state.

METHOD

Sixteen right-handed subjects performed two experiments in counterbalanced order. Both experiments used the same exposure task: inspecting the subject's own feet through a prism. This technique has been shown (Craske & Crawshaw, 1974; Crawshaw & Craske, 1974) to produce reliable prism adaptation of registered eye position alone without involving hands and arms. Both experiments consisted of an exposure condition in which a prism was used, following a control condition. Each condition consisted of eight 60 sec inspection periods alternating with measurements.

EXPERIMENT I

Procedure

The apparatus and procedure were essentially similar to those described in detail in Craske & Crawshaw (1974). Briefly, there was a horizontal board at shoulder height with a bite board on one side. Opposite, 4 microlight targets were arranged horizontally 2° apart. The subject could indicate his judgement of target direction by marking a piece of paper fixed to the underside of the board with a hand-held pen. A mask was arranged before the subject's face with a shutter, controlled remotely by the experimenter, before each eye. A pair of spectacle trial frames were used in the inspection periods. These could be fitted with a 20 dioptre base-nasal prism on one side and a diffusing occluder on the other. The frames were surrounded by an opaque cloth mask to prevent accidental undeviated vision.

In the control condition the subject inspected his feet while wearing the spectacles with the diffusing occluder before one eye but nothing before the other. After 60 sec the room lights were extinguished and the subject removed the spectacles and fitted his teeth to the bite. The experimenter opened one shutter in the mask and then switched on one of the target lights. The subject made three separate indications of the direction of the target. The target was extinguished and the shutter closed. The subject relinquished the bite and put the spectacles on once more before the room lights were turned on to begin another 60 sec inspection period.

All four target lights were presented, in random order, to each eye once in each condition.

The exposure condition was the same as the control, but with a 20 dioptre base-nasal prism before the non-occluded eye. In addition, the subject's hands were kept completely out of sight in the inspection periods.

Results

The mean of the three localizations made on each trial was used for analysis.

For all subjects, the mean adaptation for both eyes combined was 1·95° (17·23 per cent of the prism power); this was significant ($F = 15{\cdot}5$, d.f. 1, 15, $P < 0{\cdot}01$). For the exposed eye alone, mean adaptation was 1·91°, and for the occluded eye, 1·98°; this difference was not significant ($F = 0{\cdot}025$, d.f. 1, 15).

EXPERIMENT II

Procedure

The frame of the apparatus already described was used, but the shuttered mask was removed and the horizontal board was replaced by a vertical one in a fronto-parallel plane immediately before the subject's face. On the side further from the subject was a piece of paper on which the subject could mark the estimated position of the tip of his nose. Three marks were made on each trial, and there were eight trials in both conditions. Thus the measures exactly matched those of Expt. I. The eye which a given subject used and every other circumstance was identical to that of Expt. I.

Results

The mean of the three localizations made on each trial was used for analysis. There was no absolute measure of accuracy possible: the interest lay in the change between conditions. This difference was not significant ($F = 1, 36$, d.f. 1, 15) indeed, only nine subjects' data were in the adaptive direction.

DISCUSSION

It is important to note that although the hands were involved in indicating direction, the site of adaptation must be elsewhere, for the hands were never seen while the prism was worn. Nor was adaptation at the level of the neck as Expt. II shows. Recently, Cohen (1974) has reported evidence suggesting a change in registered position of the head on the shoulders under certain conditions. It may be that in the present experiments the change in position within the sagittal plane between the exposure and measurement conditions obliterated such an effect. But, whatever the reason, the procedure used here seems particularly useful for the investigation of eye-position adaptation, because of the freedom from contamination.

With mean adaptation at the occluded eye being slightly greater than with the exposed eye, the question of partial transfer does not arise: the difference, however, was very far from significant. There is no doubt that the results of Expt. I show that interocular transfer of *adapted registered eye position* is complete under the conditions used.

This paper, together with that of Craske & Crawshaw (1974), suggests that the positional information derived from the oculomotor system of the two eyes is independently processed, but linked in such a way that discordant positional information into one eye when used alone is sufficient to adjust the registered position of both. On the other hand, opposing positional discordance to each eye is a sufficient condition for the positional systems of each eye to adapt independently.

Martin Crawshaw received support from SRC.

REFERENCES

BOSSOM, J. & HAMILTON, C. R. (1963). Interocular transfer of prism-altered coordinations in split-brain monkeys. *J. comp. physiol. Psychol.* **56**, 769–774.

COHEN, M. M. (1974). Changes in auditory localization following prismatic exposure under continuous and terminal visual feedback. *Percept. mot. Skills* **38**, 1202.

CRASKE, B. & CRAWSHAW, M. (1974). Adaptive changes of opposite sign in the oculomotor systems of the two eyes. *Quart. Jl exp. Psychol.* **26**, 106–113.

CRAWSHAW, M. & CRASKE, B. (1974). No retinal component of prism adaptation. *Acta psychol.* **38**, 421–423.

EBENHOLTZ, S. M. (1970). On the relation between interocular transfer of adaptation and Hering's law of equal innervation. *Psychol. Rev.* **77**, 343–347.

FOLEY, J. E. (1970). Prism adaptation with simultaneous receipt of normal input. *Percept. Psychophys.* **8**, 393–395.

FOLEY, J. E. & MIYANISHI, K. (1969). Interocular effects in prism adaptation. *Science, N.Y.* **165**, 311–312.

HAJOS, A. (1968). Sensumotorische Koordinationsprozesse bei Richtungslokalisation. *Z. exp. angew. Psychol.* **15**, 435–461.

HAJOS, A. & RITTER, M. (1965). Experiments to the problem of interocular transfer. *Acta psychol.* **24**, 81–90.

HAMILTON, C. R. (1967). Effects of brain bisection on eye–hand coordination in monkeys wearing prisms. *J. comp. physiol. Psychol.* **64**, 434–443.

HARRIS, C. S. (1965). Perceptual adaptation to inverted, reversed, and displaced vision. *Psychol. Rev.* **72**, 419–444.

HAY, J. C. & PICK, H. L. (1966). Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *J. exp. Psychol.* **71**, 150–158.

KOHLER, I. (1963). The formation and transformation of the perceptual world. *Psychol. Issues* **3** (4).

LUND, J. S. (1965). Preliminary report of studies on the adaptation of normal and brain operated monkeys to a prism-induced deviation of the visual field. Privately circulated.

PICK, H. L., HAY, J. C. & PABST, J. (1963). Kinesthetic adaptation to visual distortion. Presented to M. P. A., May.

(*Manuscript received* 25 *January* 1975; *revised manuscript received* 2 *July* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 479–486
Printed in Great Britain

RATIONALIZATION IN A REASONING TASK

By J. St B. T. EVANS
Plymouth Polytechnic
and P. C. WASON
University College London

An experiment is presented to test the dual process model of thinking proposed by Wason & Evans (1975). One prediction of the model is that subjects' introspective justifications of their reasoning behaviour do not reflect insight into the underlying thought processes, but are essentially rationalizations. In the present experiment subjects in independent groups are given one of several different solutions to a reasoning problem as being correct, and asked to justify the solution. As predicted, subjects tend to find reasons for whichever solution is presented and express a generally high degree of confidence in the correctness of their reasons. The discussion relates both the results and the underlying theory to concepts developed by social psychologists.

Wason & Evans (1975) have proposed that an explanation of reasoning performance entails a distinction between two kinds of thought process. First, there are the processes which underlie the performance from which reasoning is inferred, e.g. the choice of a certain conclusion as valid, or the selection of potentially relevant information. We maintain that the processes which underlie such behaviour (Type 1 processes) are not generally available to introspection. The other kind of process (Type 2 process) operates at a conscious level, and hence may be reported by the subject as the reason for his performance. These Type 2 processes may have some part in organizing behaviour, but their most common function is to *rationalize* the outcome of Type 1 processes of which the subject is unaware. When the reasoning is correct, the Type 2 processes are assumed to reflect the Type 1 processes, but we argue that they should still be regarded as rationalizations of an intuitive insight or overt response. Our results have strongly suggested such a dual process model, and we have cited other results in reasoning experiments which support the necessity of a distinction between two kinds of process (see Wason & Evans, 1975).

The particular context from which the model has emerged is a deceptively simple problem devised by Wason (1966), known as the 'selection task', or 'four-card problem'. The paradigm case is as follows: a subject is shown four cards displaying B, D, 3, 7, respectively. He knows that each card has a letter on one side and a number on the other side. He is then given the following test sentence: 'If there is a B on one side of a card, then there is a 3 on the other side of the card.' The problem is to decide which cards need to be turned over in order to determine the truth value of the sentence. The solution is the B and the 7, since only a B combined with some number *other* than 3 could *falsify* the sentence in question. With abstract material most individuals fail to solve the problem; they generally omit the 7, and redundantly choose the 3 (see Wason & Johnson-Laird, 1972).

A detailed discussion of the different theoretical explanations of performance, leading up to the dual process model, is given by Evans (1974). The essential points can be summarized briefly: Wason's original suggestion, that error is due to a tendency to *verify* rather than *falsify* the test sentence, has been formalized by Johnson-Laird & Wason (1970). They operationally defined three levels of 'insight' assumed to underlie the different solutions: with 'no insight' a subject will choose only cards which could verify the test sentence; with 'partial insight', cards which could both verify and falsify; with 'complete insight', only cards which could falsify. Goodwin & Wason (1972) found that introspections corresponded to the states of insight defined by this model, thus apparently corroborating it. However, Evans & Lynch (1973) found that a factor, discovered on a different reasoning problem (Evans, 1972), could account for performance with greater parsimony. This factor, 'matching bias', is simply a tendency to select the values *named* in the test sentence (B and 3 in the above example), *regardless* of the presence of negatives in the sentence. Since the logical consequences of 'matching' are altered when a negative is introduced, the Johnson-Laird & Wason (1970) model can account for performance with reference *only* to affirmative test sentences.

If matching bias underlies the selection of cards, then why do the introspections reported by Goodwin & Wason (1972) indicate 'degrees of insight' in terms of verification and falsification? According to the dual process model, matching bias is a Type 1 process of which the subject is unaware. When asked to justify behaviour, however, the Type 2 processes construct a rationalization of performance in the context set by the instructions, i.e. 'to test whether the sentence is true or false'. The small proportion of subjects (18·75 per cent) who attained 'complete insight', i.e. solved the problem, are assumed to have overcome the effects of matching bias, and correctly appreciated the relevance of falsification. The dual process model generates different predictions from the Johnson-Laird & Wason (1970) model only with respect to error: it substitutes 'matching' for 'verification' as the most frequent process underlying attempts to solve the problem.

Wason & Evans (1975) tested the hypothesis that introspections in the task are generally constructive rationalizations. They compared the justifications of performance under a test sentence in which the consequent is negated with those under a test sentence in which it is not negated. When the consequent is negated matching bias would induce a spuriously correct solution; when it is not negated, it would induce a spuriously incorrect solution. For example, given the test sentence 'if D . . . then not 4', the correct solution is D and 4; given the sentence 'if D . . . then 4', the correct solution is D and a number other than 4. The results showed that the selections were primarily consistent with the effects of matching bias on both types of test sentence, but the justifications were strikingly different. There was a strong tendency for the subjects to state that they were trying to falsify the negated test sentence, thus apparently revealing 'complete insight' into the structure of the problem. But the same subjects revealed a total lack of such insight with respect to the affirmative test sentence.

Wason & Evans' (1975) results do appear to support the idea that these justifications are rationalizations, but a further experimental test of the hypothesis is obviously needed. While recognizing the need to test the dual process model in different situations, the present study involves the selection task in order to reconsider some previous results obtained in a pilot study (Wason, 1969). The experiment was very simple: the subjects were first given the correct solution, and *then* asked to explain it.

All 20 subjects did so correctly. At the time Wason supposed that the subjects had been prevented from imposing their own erroneous structure on the problem, but (strictly) this supposition implies that the subjects would also appreciate the erroneous nature of a given wrong solution, and this condition was not investigated. Indeed, in the light of the present hypothesis an alternative, more parsimonious interpretation is that the subjects had been induced to produce a Type 2 thought process, effectively rationalizing someone else's solution.

In the present experiment the subjects are asked to justify three common erroneous solutions and the correct solution, all of which purport to be the *correct solution*. It was predicted that the subjects would not only tend to give reasons supporting any given solution, but also indicate a high degree of confidence in the correctness of their justifications.

METHOD

Design

Four independent groups were used, each being presented with a different solution to a selection task. The test sentence had the following general form: 'If there is L+ on one side of a card, then there is N+ on the other side of the card.' The symbol L+ refers to the letter named in the antecedent of the sentence (and to the corresponding card), L− refers to an unnamed letter (and corresponding card), N+ to the number named in the consequent (and corresponding card), and N− to an unnamed number (and corresponding card). The solutions given to the four groups were L+N− (correct), L+N−N+ (partially correct), L+N+ and L+ (both incorrect).

Subjects

Forty first year students on a Business Studies degree at Plymouth Polytechnic were randomly allocated to four experimental conditions in equal proportions, and tested as a single group.

Task and instructions

The instructions and problems for each subject were laid out on two duplicated sheets, the first being identical for all groups. The instructions began as follows: 'This experiment is designed to examine the extent to which people have an intuitive understanding of logical reasoning without the benefit of formal training. If you have received training in formal logic, please inform the experimenter. The problem is concerned with a set of four cards lying on a table: naturally only one side of each card is visible. The following information is relevant to the problem and must be kept in mind: *Every card has a letter on one side and a number on the other side.*'

Underneath was a diagram of four cards showing the symbols, D, B, 3, and 7. Beneath there was presented the following rule: 'If there is a D on one side of the card, then there is a 3 on the other side of the card.'

The subject was instructed that the rule applied only to these four cards and that it might be true or false.

The following instructions were given on the second sheet, differing between groups only in respect to the solution:

'*The logical problem*: On the basis of the information given on the first sheet, you should now consider a logical problem which you are *not* required to solve; indeed the solution is given here. The task you have to perform is explained below. The problem is to decide, on the basis of the information given, which of the four cards need necessarily be turned over in order to find out whether the rule is true or false.

'*Solution*: It is necessary to turn over the D and the 7. It is not necessary to turn over the B or the 3.

'*Your task*: What you are required to do is to explain the reason why it is necessary or unnecessary to turn over each card. Please write in the reasons (one for each card) in the appropriate places below. It would be helpful if you could also express your degree of confidence in each reason you give. We want to know how sure you are that each reason you have given is the correct one. Please do this by entering a number between 1 and 4 in the column headed "confidence" corresponding to the following scale: 1 = highly confident 2 = confident 3 = not very confident 4 = not at all confident.'

The instructions given to the different groups varied only with respect to the solution: the above example is for group L+N−. Group L+N+N− was told that D, 3 and 7 were necessary, and B unnecessary. Similarly, group L+N+ was told that D and 3 were necessary, and B and 7 unnecessary; and group L+ was told that D was necessary and B, 3 and 7 unnecessary. Below these instructions appropriate spaces were provided for the subjects to state their reasons for selecting, or rejecting, each card, together with the confidence rating for each reason. No time limit was imposed on the task.

Results

An attempt was first made to determine the extent of agreement between the given solutions and the reasons elicited from the subjects. The two authors independently rated each subject's response on each card with respect to agreement or disagreement with the solution. A comparison of the two ratings revealed very few inconsistencies, and these were easily resolved. In order to make a stronger test of the rationalization hypothesis, a classification of 'confident agreement' (CA) was made. The criterion adopted was that a stated reason should both be rated as agreeing with the solution and assigned a confidence level 1 or 2 by the subject.

A global analysis was carried out first of all, categorizing the subjects according to whether they had consistent confident agreement, i.e. a CA classification *on all four cards*. The number of subjects thus classified was 3 out of 10 for both group L+N− and group L+N+N−. These solutions are correct and partially correct (respectively) and they occur infrequently with abstract material (Wason & Johnson-Laird, 1972). On the other hand, the number thus classified for group L+N+ was 6 out of 10 and for group L+, 7 out of 10. These solutions are both incorrect, and occur very frequently with abstract material. Thus, consistent confident agreement was observed for only 6 out of 20 subjects given the rare (and correct) solutions, but for 13 out of 20 subjects given the frequent (and incorrect) solutions. This result is consistent with a *single* process model, according to which a subject would be more likely to agree with those solutions which he would be most likely to make under standard conditions of presentation. The trend, is, in fact, just below statistical significance ($\chi^2 = 3{\cdot}61$).

Table 1 shows the corresponding analysis for *each* card in relation to the given solutions. It gives the frequency of agreements (at both confidence levels), and the disagreements with respect to selection or omission. In all four groups, the subjects were told to select L+ and not to select L−, but instructions concerning N+ and N− varied. If these two critical cards are considered alone, it will be noted that the total frequency of CA responses for them is 16 in the correct solution, L+N−. And this is close to the corresponding frequency in the two erroneous, frequent solutions, L+N+ and L+ i.e. 17 and 16 respectively. In fact, it is the corresponding frequency of CA responses under the partially correct solution, L+N+N−, which is somewhat depressed, i.e. 11.

Inspection of Table 1, however, still supports the single process hypothesis, with the qualification that the frequency for partially correct confident agreements is anomalous with respect to previous results. But a closer inspection shows that this

Table 1. *Degree of agreement with the given solution for each card*

	Groups											
	L+N+			L+			L+N+N−			L+N−		
	CA	UA	D	CA	UA	D	CA	UA	D	CA	UA	D
Cards												
L+	9	1	0	10	0	0	7	3	0	7	2	1
L−	8	0	2	8	1	1	6	2	2	7	2	1
N+	10	0	0	8	1	1	5	4	1	8	2	0
N−	7	1	2	8	1	1	6	2	2	8	2	0
Total	34	2	5	34	3	3	24	11	5	30	8	2

CA = confident agreement; UA = unconfident agreement; D = disagreement.

Table 2. *Confident agreement with instructions to select, or not select, the two critical cards, N+ and N− (maximum possible frequency for each cell equals* 20)

	Instruction	
	Select	Not select
Cards		
N+	15	16
N−	14	15

result is due at least as much to CA responses on the *unvaried* L+ and L− cards, as to CA responses on the *critical* N+ and N− cards. It is, accordingly, imperative to examine the CA responses to the two critical cards N+ and N−, irrespective of the particular solutions in which they are assessed. Table 2 shows these frequencies.

It will be observed that in Table 2 there is virtually no difference in the frequency of CA responses to N+ and N− regardless of whether the solution specified their selection or non-selection. And this result is consistent with the *dual* process model.

Discussion

The results reveal an apparently puzzling inconsistency. The global analysis of confident agreement with the *solutions* fits a *single* process model because the subjects agree more readily with their reasons for solutions which they would have been most likely to make under standard conditions of presentation. On the other hand, an individual *card* analysis fits the *dual* process model because there is virtually no difference between the frequencies with which subjects confidently agree about their reasons about N+ (31) and N− (29), irrespective of whether the solution specifies selection or non-selection of these two critical cards. The statistical reason for this result can be seen by inspecting Table 1. In the infrequent solutions, those specifying the selection of N−, confidence is decreased in the reasons for all the cards. If this were due to generalization from a disinclination to agree confidently with the selection of N−, it would be expected that such lack of agreement would be most pronounced with respect to N−. But (if anything) almost the opposite occurred. The reasons associated with the two correct cards, L+ and N−, are *least* affected in their confidence rating.

This observation resolves the apparent paradox of the two analyses but leaves a difficult psychological problem. If the unusual, and infrequent, selection of N − had facilitated any insight into the solution, it would be expected that confident agreement would *increase* for reasons associated with the (correct) non-selection of L − and N +, as well as for the (correct) selection of L + and N −. The fact that this did not occur is evidence against the original interpretation of Wason's (1969) results. In that experiment it was inferred that subjects were correct because they had not been allowed to impose their own erroneous structure on the problem. The present results suggest that the presentation of the correct solution does not induce enlightenment.

In the present experiment, of the 20 subjects who were told the solution involved the selection of N −, 14 indicated confident agreement with their reasons for solution, and 11 of these offered reasons which suggested apparent insight into the relevance of falsification, e.g. 'To see whether or not a D (L +) is present on the reverse side. If it is then the statement is obviously false', 'If a D (L +) was on the reverse, then the rule would be proved false', 'A D (L +) could also be on this last card, for the proof does not state that the opposite side is restricted to 3 (N +)', 'Must turn over or you would not know whether the rule applies or not'.

In contrast, of the 20 subjects who were told the solution did not involve the selection of N −, only three disagreed. One commented 'If you do not turn over this you will not know if there is a D (L +) on the other side and cannot know if the solution is correct'. The most typical reason, of the 17 who agreed with the non-selection of N −, was 'Because the number is 7 (N −), and therefore the letter on the other side is irrelevant'.

In some cases the reasons given strongly suggested a rationalizing process. This is exemplified by the following protocol of a subject given L + N + N − as the solution. (Our symbols have been substituted for the actual letters and numbers.)

Card	Turn over?	Reasons	Confidence
L +	Yes	The first part of the rule is obeyed so it is necessary to turn over to see if all the rule is true	1
L −	No	The rule is not obeyed to start with so there is no point in turning over	1
N +	Yes	The second part of the rule is obeyed so it is necessary to turn over to see if the first part is as well	1
N −	Yes	The second part is not obeyed, but you don't know if the first part is so you must turn the card over or you would not know whether the rule applies or not	1

Consider first the reasons given to justify the role of the first three cards in the solution. They all accord with matching. The named letter (L +) and number (N +) are selected, but not the unnamed letter (L −). The evidence for the operation of a matching response bias is revealed by the subject's emphasis that the rule is 'obeyed', or 'not obeyed'. Furthermore, we see from his reasons for L + and N + that his justification for their selection is based on finding a verifying combination. Here then we have an apparent matcher/verifier who, if solving the problem, would presumably have rejected N–. However, while observing that it mismatches (is 'not obeyed') he appears to convince himself by rather tortuous reasoning that its selection is necessary. It is interesting to note that, as with the previous reasons, he does so with maximal confidence.

On balance, then, the results of this experiment do support the rationalization hypothesis of the dual process model. The critical reader, however, might object that what we have done is essentially an experiment in social psychology, inducing subjects to comply with an instruction reinforced by the authority of the experimenter. Certainly a strong set is induced for the subject to believe that a given solution is correct, and in justifying it he may be responding to social pressure. However, if the dual process model is correct, this is not different *in principle* from such experiments as Goodwin & Wason (1972) and Wason & Evans (1975), in which the subject is encouraged to give his own solution to the problem. The request to explain a thought process, of which the subject is unaware, may also be construed as a form of social pressure. And the subject's willingness to produce such an 'explanation' could be regarded as an instance of 'cognitive consistency' (Zajonc, 1968). As social psychologists have shown, individuals like to maintain a consistent picture of their own behaviour in relation to their current belief system. However, such a reorganization of attitudes need involve only Type 2 thought processes rather than Type 1 processes. Hence the behavioural implications of much research into attitude change may be considerably less than some social psychologists assume.

If we regard our subjects' rationalizations as a form of attitude change, than we should also consider the nature of the change. Attitude theorists distinguish between 'compliance', in which the subject merely conforms in his behaviour, e.g. the classic study of Asch (1956), and 'internalization', indicative of a genuine change in the subject's belief (see Kelman, 1961). The use of confidence ratings in the present experiments helps to distinguish between these two contingencies. If the subjects are giving reasons to support a given solution, *merely as a form of compliance*, then their confidence for such reasons should be low. Compliance does not entail high confidence in the reasons for a given solution because the subject is aware that his degree of confidence for a reason, offered for any selection, may vary from one to four. But the fact that a high degree of confidence was generally given suggests that belief in the solution was genuinely 'internalized'. Evidence for such 'internalization' in the standard presentation of the task, revealed by deeply fixated responses, is abundantly illustrated in previous research on the problem (Wason & Johnson-Laird, 1972).

It is relevant here to make some general remarks about the attitude towards the dual process model. We have sensed that some of our colleagues, who have worked on the selection task, are resistant to it, perhaps because it seems to impugn rationality. One exclaimed: 'It *can't* be right'.

In fact, the hypothesis does not impugn rationality any more than does the pervasiveness of intuitive thought or the phenomenon of post-hypnotic suggestion. However, the temporal relation between 'selection' and 'justification' is complex. We have suggested that, according to the common-sense *rational hypothesis*, the 'reasons' (justifications) temporally precede the 'selection'. The strong form of the dual process hypothesis postulates the opposite: the selection temporally precedes the reason.

A weaker version is that Type 1 and 2 processes are concurrent, and interact with rapid feedback between them (Wason & Evans, 1975).

The strong version seems counterintuitive, but there is one striking analogy to it. It is a fairly common experience to wake up with a convulsive jerk of the entire musculature which coincides with an 'appropriate' event in a dream sequence, e.g. stumbling down a step. In recall, the ultimate event seems a plausible culmination of the dream: 'I was walking down some narrow steps, and *then* I stumbled'. This is puzzling because it is as if the dream sequence had anticipated (or even caused) the subsequent convulsive jerk. The impression is almost certainly an illusion. The jerk had broken sleep, initiating a rationalizing dream with an illusory temporal sequence. And this is like the strong version of the dual process hypothesis: a response precedes a rationalization, even though we may have the illusion that a 'reason' precedes the response.

The authors would like to thank Dr Stephen Newstead for suggesting an improvement in the design and assisting with the collection of the data.

References

Asch, S. E. (1956). Studies of independence and conformity. *Psychol. Monogr.* **70**, no. 9. (Whole no. 416.)

Evans, J. St B. T. (1972). Interpretation and matching bias in a reasoning task. *Q. Jl exp. Psychol.* **24**, 193–199.

Evans, J. St B. T. (1974). On the origin of selections in the selection task. Paper read to the Trento conference on the selection task.

Evans, J. St B. T. & Lynch, J. S. (1973). Matching bias in the selection task. *Br. J. Psychol.* **64**, 391–397.

Goodwin, R. Q. & Wason, P. C. (1972). Degrees of insight. *Br. J. Psychol.* **63**, 205–212.

Johnson-Laird, P. N. & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cog. Psychol.* **1**, 134–138.

Kelman, H. C. (1961). Processes of opinion change. *Publ. Opin. Q.* **25**, 57–58.

Wason, P. C. (1966). Reasoning. In B. M. Foss (ed.), *New Horizons in Psychology* I. Harmondsworth: Penguin.

Wason, P. C. (1969). Structural simplicity and psychological complexity: some thoughts on a novel problem. *Bull. Br. psychol. Soc.* **22**, 281–284.

Wason, P. C. & Evans, J. St B. T. (1975). Dual processes in reasoning? *Cognition* **3** (in press).

Wason, P. C. & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content.* London: Batsford.

Zajonc, A. B. (1968). Cognitive theories in social psychology. In G. Lindsey & E. Aronson (eds), *The Handbook of Social Psychology*, vol. I. London: Addison-Wesley.

(*Manuscript received* 18 *July* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 487–500
Printed in Great Britain

PROCEDURES FOR INVESTIGATING IMAGERY AND THE DISTINCTION BETWEEN PRIMARY AND SECONDARY MEMORY

BY JOHN T. E. RICHARDSON*

Neuropsychology Unit, Department of Neurology, The Churchill Hospital, Oxford

Previous research has shown that the imageability of the material affects the secondary memory (SM) component of free recall, but not the primary memory (PM) component. However, other research has shown that imageability affects the recall of items of low concreteness, but not of those of high concreteness. Expt. I related these findings by showing that the effect of imageability on items of low concreteness applied to SM, but not to PM. However, the effect of the imagery ability of the subjects did not accord well with this framework. Subjects of high imagery ability produced a larger recency effect, but a smaller primacy effect. It was suggested that subjects of high imagery ability use imagery in learning, but that they fail to use it in an interactive way. This was supported by the results of Expt. II, which found that subjects given separative imagery instructions produced similar serial position curves to subjects of high imagery ability. On the other hand, interactive imagery instructions produced an increase in performance over all serial positions. It was concluded that separative imagery instructions disrupt performance by interfering with the subject's organization of a list, and that the PM/SM distinction should be extended to take into account the variety of possible strategies employed by subjects in learning.

There now exists considerable support for the proposition that imagery is an important factor in many psychological tasks, and it has recently been suggested (Richardson, 1974) that it would be useful to relate imagery to other theoretical concepts, in particular, the distinction between primary memory (PM) and secondary memory (SM). Waugh & Norman (1965) proposed that, in the free recall of a list of unrelated words, the last few items are retrieved easily from a short-term store of limited capacity (PM), whereas the remaining items are retrieved with some difficulty from a long-term store of large capacity (SM). Consequently, the SM component of free recall is represented in the primacy and middle portions of the serial position curve, and the PM component is represented in the recency portion (Glanzer & Cunitz, 1966). It has been shown that the positive effect of the imageability of stimulus material upon recall performance applies to the SM component, but not to the PM component (Richardson, 1974). Moreover, in a paired associate learning task, Smith, Barresi & Gross (1971) found that the positive effect of instructions to use imagery in learning was due to a facilitation of SM, but not of PM. The first experiment to be reported here extends this research by considering the role of individual differences in imagery ability.

The ability to use mental imagery is assessed either by ratings of the subjective vividness of experienced imagery, as in Betts' (1909) Questionnaire Upon Mental Imagery and in Sheehan's (1967*a*) shortened version, or by tests of spatial thinking, such as the Flags Test (Thurstone & Jeffrey, 1956). Factor analytic studies have indicated that these two procedures define independent attributes (Di Vesta, Ingersoll & Sunshine, 1971; Paivio, 1971, pp. 495–497); indeed, the study by Di Vesta

---

* Present address: Department of Psychology, Brunel University, Uxbridge, Middlesex UB8 3PH.

et al. found that imagery, as measured by introspective reports (including Sheehan's test), was saturated on a factor of social desirability, and they concluded that introspective reports did not possess construct validity as measures of imagery ability. Indeed, early evidence on the function of individual differences in memory tasks derived from the use of introspective methods was generally contradictory or negative (Paivio, 1971, p. 503). Nevertheless, Marks (1972) reported a positive relationship between subjectively rated imagery ability and performance in a picture-memory task, while Sheehan found such a relationship in some studies (Sheehan, 1966, 1967*b*), but not in others (Sheehan & Neisser, 1969).

On the other hand, experiments using tests of spatial thinking suggest that subjects of high spatial ability perform better than subjects of low spatial ability on pictorial material and in tasks favouring non-verbal mediation, whereas the reverse is true on abstract material and in tasks favouring verbal mediation (Hollenberg, 1970; Stewart, 1965, described in Paivio, 1971, pp. 505–508). Using a recognition task for words and pictures, Stewart found results suggesting that subjects of high spatial ability were more likely to encode a word as a picture than were those of low spatial ability, whereas subjects of low spatial ability were more likely to encode a picture into a word than were those of high spatial ability. Unfortunately, she failed to find a similar interaction between spatial ability and the vividness of the material in an experiment involving the free recall of lists of words, and Di Vesta *et al.* (1971) remarked that they had been 'only partially successful' in attempting to replicate Stewart's findings.

The evidence concerning the role of individual differences in imagery in memory tasks is therefore somewhat equivocal. One problem underlying the research is a degree of confusion over what the tests employed actually measure. Tests of spatial thinking are explicitly tests of some intellectual ability. However, it is not clear whether subjective questionnaires should be interpreted also as measuring some ability, or as measuring the subject's preference for using imagery as opposed to some other strategy. This ambiguity over the interpretation of introspective reports makes it more difficult to offer clear predictions concerning the effects of individual differences upon performance, and this point, together with the conclusions of Di Vesta *et al.* (1971) tends to favour the use of objective tests of spatial thinking. The experiment reported below employed such a test to investigate the effect of imagery ability upon performance in free recall, and to relate such an effect to the distinction between PM and SM.

However, one must reconsider the effect of stimulus imageability for a moment. Recent experiments have shown this effect to be complicated by the confounded variable of stimulus concreteness. Like imageability, this measure is derived from experimental tasks where large numbers of subjects are asked to rate items along a seven-point scale (Paivio, Yuille & Madigan, 1968). The instructions define imageability as the extent to which the item evokes a mental image, and they define concreteness as the extent to which its referent can be experienced by the senses. Richardson (1975*a*) manipulated the two variables independently in free recall, and found that imageability had a positive effect upon the recall of items of low concreteness, but no effect upon the recall of items of high concreteness. In a subsequent experiment, it was found that imageability would affect the recall of items of high concreteness if the subjects were told to use imagery in their memorizing (Richardson, 1975*b*). This suggested that imagery is used to remember items of low concreteness, but that it is not usually employed to remember items of high concreteness. On this account, the imagery ability of the subject should affect his performance on items of low concreteness, but not on those of high concreteness.

EXPERIMENT I

Method

Subjects

This study employed 80 students from the University of Sussex and the Brighton College of Education, who were paid for their participation.

Materials

The stimuli were those of Expt. II of Richardson (1975*a*). They comprise eight lists of ten nouns, varying orthogonally rated imageability, rated concreteness, and simple versus derived nouns. The material was matched for frequency (Thorndike & Lorge, 1944), and the selection of items from Paivio *et al.* (1968) was validated by *post hoc* rating studies with English subjects.

Procedure

The material was presented to ten groups of eight subjects in homogeneous lists for immediate free recall. The order of presentation of the lists to each subject in each group was determined by an 8 x 8 Latin square, and the order of presentation of the ten items in a given list to each group of subjects was determined by a 10 x 10 Latin square. Subjects were given standard free recall instructions, words were read out by the experimenter at a rate of one item every 5 sec, and 40 sec were allowed for free oral recall.

Following the administration of the free recall task, each subject completed the Flags Test (Thurstone & Jeffrey, 1956). This contains 21 questions, each of which comprises one test figure and six examples, each being the schematic representation of a flag. The subject has to say whether each example represents the same side of the flag as the test stimulus or the opposite side, and he has to complete as many questions as possible in 5 min. The subjects were tested individually.

Results

Each subject was allocated a score on imagery ability as the difference between the number of correct responses (mean = 61·9) and the number of errors (mean = 3·2) produced in the Flags Test. The 80 subjects were divided into four sets on the basis of their rank imagery score within a group; that is, the two subjects with lowest scores were allocated to one set, the next two to another, and so on. These four sets of 20 subjects will be referred to as low, medium-low, medium-high, and high imagers. The statistical analyses to be described found no significant differences between the medium-low and medium-high imagers, and so these 40 subjects will be referred to as medium imagers. The experimental design did not balance between the order of lists and the imagery ability of a subject, but the unsystematic relationship between these two variables should reduce the possibility of such confounding to a minimum.

The mean scores for three levels of imagery ability on the four critical types of stimulus are shown in Table 1. An analysis of variance was carried out on each subject's score on each item; this employed the appropriate procedures for generalizing over both subjects and material (cf. Clark, 1973). The variation between the four types of stimulus shown was partitioned according to the *a priori* hypotheses into

Table 1. *Results of Expt. I: mean per cent correct in immediate free recall, by concreteness* (C) *and imageability* (I) *of stimulus material, and by imagery ability of subjects*

Imagery ability	High C		Low C		High I – Low I	
	High I	Low I	High I	Low I	High C	Low C
High	58·8	59·5	66·3	50·8	−0·7	15·5
Medium	56·5	52·8	57·4	51·5	3·8	5·9
Low	60·3	55·8	64·3	56·0	4·5	8·3
Overall	58·0	55·2	61·3	52·4	2·8	8·9

an effect of concreteness, an effect of imageability in the case of items of high concreteness, and an effect of imageability in the case of items of low concreteness. Previous theoretical discussions (e.g. Stewart, 1965) suggest that the appropriate comparison among the four sets of subjects is that between high imagers and the rest. The remaining variation was partitioned into the difference between low and medium imagers, and the difference between the two groups of medium imagers.

The analysis yielded a significant effect of imageability in the case of items of low concreteness ($F = 5·75$, d.f. $= 1, 72$, $P < 0·02$), and a significant interaction between high imagers and the other three sets of subjects ($F = 5·73$, d.f. $= 1, 72$, $P < 0·02$). The effect of imageability in the case of items of high concreteness and the corresponding interaction were not significant ($F < 1$, and $F = 1·86$, d.f. $= 1, 72$, $P > 0·1$). Tests on simple main effects showed that the effect of imageability in the case of items of low concreteness was significant both for the high imagers ($F = 7·00$, d.f. $= 1, 144$, $P < 0·01$), and for the other three sets combined ($F = 4·49$, d.f. $= 1, 144$, $P < 0·05$).

The analysis also yielded a significant effect of serial position ($F = 53·61$, d.f. $= 9, 576$, $P < 0·001$), and a significant interaction between this effect and the difference between low and medium imagers ($F = 2·95$, d.f. $= 9, 576$, $P < 0·01$). There were no further significant terms in the analysis of variance. The serial position curves for three levels of imagery ability are shown in Fig. 1. The effect of serial position was examined in more detail by the use of pairwise *a posteriori* tests between the three sets of subjects on the primacy effect, the middle portion, and the recency effect. The mean per cent correct for each set over serial positions 1–3 was: high imagers, 59·0; medium imagers, 53·5; and low imagers, 67·7. The low imagers were significantly different from the high imagers ($F = 5·69$, d.f. $= 1, 400$, $P < 0·02$) and from the medium imagers ($F = 19·88$, d.f. $= 1, 400$, $P < 0·001$), but the high imagers did not differ from the medium imagers ($F = 2·91$, d.f. $= 1, 400$, $P > 0·05$). The mean per cent correct for each set over serial positions 4–7 was: high imagers, 47·5; medium imagers, 44·0; and low imagers, 47·0. There were no significant differences. Finally, the mean per cent correct for each set over serial positions 8–10 was: high imagers, 73·8; medium imagers, 69·5; and low imagers, 66·5. The high imagers differed significantly from the low imagers ($F = 3·95$, d.f. $= 1, 400$, $P < 0·05$), but not from the medium imagers ($F = 1·81$, d.f. $= 1, 400$, $P > 0·1$); the medium imagers did not differ significantly from the low imagers ($F < 1$).

Waugh & Norman (1965) specified a procedure for calculating the PM component

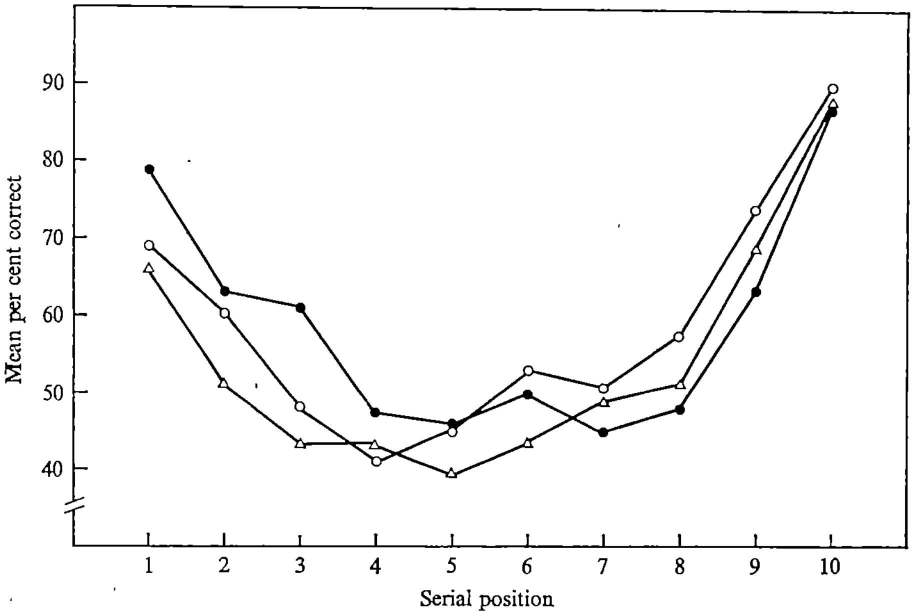


Fig. 1. Results of Expt. I: mean per cent correct in immediate free recall, by serial position and imagery ability (○—○, high imagers; △—△, medium imagers; ●—●, low imagers).

of free recall. They used the middle portion of the serial position curve to estimate the probability of an item's being retrieved from SM, and corrected the performance on the last few items for this probability, under the assumption that it was independent of the probability of an item's being retrieved from PM. The PM component for each subject on each list was calculated from the present data using serial positions 8–10. An analysis of variance was carried out on the results of these calculations, using two random factors of subjects and lists. The first was nested under the factor of imagery ability, which was partitioned as above; the second was nested under the factors of imageability and concreteness. The mean PM component for each set of subjects was: high imagers, 1·43; medium imagers, 1·08; and low imagers, 0·91. The only effect to even approach significance was the difference between high imagers and the remaining subjects ($F = 2{\cdot}65$, d.f. = 1, 76, $0{\cdot}2 > P > 0{\cdot}1$). The effect of imageability failed to be significant for either nouns of high concreteness or those of low concreteness, and the main effect of concreteness also failed to be significant ($F < 1$ in all three cases).

Finally, *a posteriori* tests were carried out within the original analysis of variance to examine the probability of an item's being retrieved from SM between the various lists. Combining scores over serial positions 4–7, it was found that the effect of imageability was significant in the case of items of low concreteness ($F = 10{\cdot}29$, d.f. = 1, 648, $P < 0{\cdot}005$), but not in the case of items of high concreteness ($F = 2{\cdot}21$, d.f. = 1, 648, $P > 0{\cdot}1$). The main effect of concreteness was not significant ($F < 1$).

Discussion

This experiment has reproduced the findings of Richardson (1975*a*), that imageability predicts the recall of items of low concreteness, but not that of items of high concreteness. Further, high imagers produced a significantly greater effect of imageability on the items of low concreteness than the medium and low imagers. This supports the idea that an effect of rated imageability indicates the use of imaginal processes by the subjects, and it constitutes evidence linking two logically distinct procedures for investigating imagery. If one considers also the results of Richardson (1975*b*), there now exists empirical foundation for the position adopted by Paivio (1971), that the image-evoking quality of stimulus material, the spatial ability of the subjects, and instructions to the subjects to use imagery in learning should be related to a single theoretical concept of imagery. On the other hand, the results of Richardson (1975*b*) suggested that imagery was only used to remember the small number of items which were of low concreteness but of high imageability, unless the subjects were explicitly instructed to use imagery. If this were so, it would explain the small and inconsistent effects of individual differences in imagery ability which previous research has produced. It also leads to a prediction for future research, that such effects will become clearer if subjects are instructed to use imagery in their learning.

An earlier study showed that the imageability of the material affected the SM component of free recall, but not the PM component (Richardson, 1974). This research confounded the variables of imageability and concreteness, but the present experiment shows that the hypothesis holds even when the two variables are manipulated independently. The effect of imageability on the recall of items of low concreteness did not interact with serial position ($F < 1$), it did not affect the PM component, as calculated according to the procedure given by Waugh & Norman (1965), but it did not affect the SM component. On the other hand, there was no effect of imageability on the recall of items of high concreteness, and this term did not affect either the PM component or the SM component of free recall. Therefore, this study has not yielded any results to contradict the position that imageability is a factor influencing secondary memory.

If one attempts to use the present results to relate imagery ability to the PM/SM distinction, however, the picture is much less clear. The effect of imagery ability upon recall performance was found to interact with serial position, such that low imagers produced a greater primacy effect, while high imagers produced a greater recency effect. Moreover, there was no effect of imagery ability upon the middle portion of the serial position curve. These results do not accord well with the approach of Glanzer & Cunitz (1966), who attempted to identify factors which affected only the primacy and middle portions, and those which affected only the recency portion of the serial position curve. Nor do they accord well with the approach of Waugh & Norman (1965), who excluded the primary effect in calculating effects on PM and SM. In terms of their model, there was no effect of imagery ability upon the probability of an item's being retrieved from SM, calculated as the average performance over serial positions 4–7. On the other hand, there was little sound evidence for the hypothesis of an effect upon the probability of an item's being retrieved from PM, although there was a trend in this direction.

A different approach is that of Rundus (1971), who associated the primacy effect in free recall with the use of a short-term store as a working memory to rehearse and recode the material, and the recency effect with the use of the same store as a temporary register. One might interpret the present results as showing that there is a trade-off between these two functions of the short-term store, varying with the subjects's imagery ability. An alternative conception is that of Baddeley & Hitch (1974), who suggested that a short-term working memory was involved in free recall, but that this store was not responsible for the recency effect. They regarded the latter as the result of an ordinal retrieval strategy (Tulving, 1968), which might be applied to information in either long-term or short-term stores. This sort of account might interpret the present findings as illustrating a trade-off between two kinds of retrieval strategy. Finally, one might relate the results of Expt. I to those of Craik (1969), who compared the effect of instructions to recall from the beginning of a list with that of instructions to recall from the end. Subjects who were told to recall from the beginning of a list produced a larger primacy effect and a smaller recency effect than those who were told to recall from the end of the list. The interaction between serial position and imagery ability in the present study might therefore reflect differences in the order of recall between the groups of subjects.

None of these last three accounts would be able to offer a complete explanation of the present results, unless some reason could be found why a subject's spatial ability should predict the extent to which he employs his short-term store as a register, the extent to which he employs an ordinal retrieval strategy to recall the last few items presented, or the extent to which he recalls from the end of a list, rather than the beginning. Nevertheless, these approaches would agree in attempting to explain the results in terms of coding or retrieval strategies. It might therefore be the case that subjects of high spatial ability use strategies based upon imagery, but not ones which lead to substantial improvements in performance; and it would be both interesting and appropriate to compare the present findings with what actually happens when subjects are given explicit instructions to use strategies based upon imagery.

EXPERIMENT II

Paivio (1971, chap. 10) has given an excellent account of the literature concerning the effects of instructions to use imagery, and it will be sufficient to summarize the main findings here. Instructions to employ imagery in learning usually produce large improvements in performance, and techniques which involve the use of imagery can be powerful mnemonic aids. However, instructions to produce images for individual items may have no effect on performance (Bower, 1970, 1972), and it would appear that where imagery is helpful it has an effect by increasing the organization and cohesion of the material which is to be remembered. In a paired-associate learning task, Smith *et al.* (1971) found that the effect applied to SM but not to PM, and the following experiment examined this hypothesis in the case of free recall.

As mentioned above, it was found in an earlier experiment that the interaction between imageability and concreteness in their effects upon free recall performance disappeared when subjects were instructed to use imagery in learning (Richardson, 1975*b*). This was so whether they were instructed to construct images for individual items (separative instructions), or whether they were instructed to construct images linking successive items in a list (interactive instructions). It was concluded that subjects given separative instructions do use imagery in their learning, even though this does not produce any increase in performance. Referring to the discussion of the previous experiment, one might suggest that subjects of high imagery ability use the same strategies in learning as those given separative instructions. Consequently, subjects who are given such instructions should produce the same kind of relationship between performance and serial position as those of high spatial ability.

Method

Subjects

This experiment employed 64 students from the University of Sussex and the Brighton College of Education, who were paid for their participation.

Materials

A set of 160 nouns was constructed using non-homophonic, English nouns with a frequency of AA or A in the count of Thorndike & Lorge (1944). These were randomly allocated to 16 lists, each of ten words. The lists were allocated at random to two blocks, which will be referred to as Blocks A and B.

Procedure

Four groups of 16 subjects were tested. For their first block of lists, all subjects were given standard free recall instructions; this is referred to as Condition I. Following this, three groups were given additional instructions for their second blocks of lists; this is referred to as Condition II. All of these groups were told that 'experiments have shown that a useful method of trying to remember a list of words is to try to make up a mental image of the thing described by each word'. Group 1 were told to make up such an image for each word in the subsequent lists. Group 2 were told that it would be particularly helpful to make up images relating successive pairs of words. Group 3 were told that it would be particularly helpful to make up images relating consecutive groups of words, e.g. one relating the first, second and third words, another relating the second, third and fourth, etc. Group 0 were not given any further instructions before the second block of lists.

The words were presented on plain white cards, typed in lower-case letters. After each list, a card bearing a row of asterisks was the cue for immediate written recall. The words were presented at a rate of one every 5 sec, and 40 sec were allowed for recall. A break of approximately 1 min was taken between the two blocks of lists.

The subjects were tested individually. Half of those in each group received Block A, followed by Block B, and half received Block B, followed by Block A. The order of presentation of the lists within each block to the subjects in each group-half was determined by an 8 × 8 Latin square. The order of presentation of the words in each list was constant.

Results

Fig. 2 compares each of the four groups under Conditions I and II. An analysis of variance was carried out on each subject's total correct responses under each condition at each serial position. The variation between the four groups of subjects was partitioned into three components: the difference between Group 0 and Groups 1–3, i.e. the effect of imagery instructions; the difference between Group 1 and Groups 2 and 3, i.e. separative vs. interactive instructions; and the difference between Group 2 and Group 3.

The analysis produced a significant difference between Conditions I and II ($F = 12{\cdot}01$, d.f. = 1, 60, $P < 0{\cdot}001$), but this was qualified by an interaction with the difference between Group 1 and Groups 2 and 3 ($F = 20{\cdot}14$, d.f. = 1, 60, $P < 0{\cdot}001$). There was a significant effect of serial position ($F = 15{\cdot}19$, d.f. = 9,

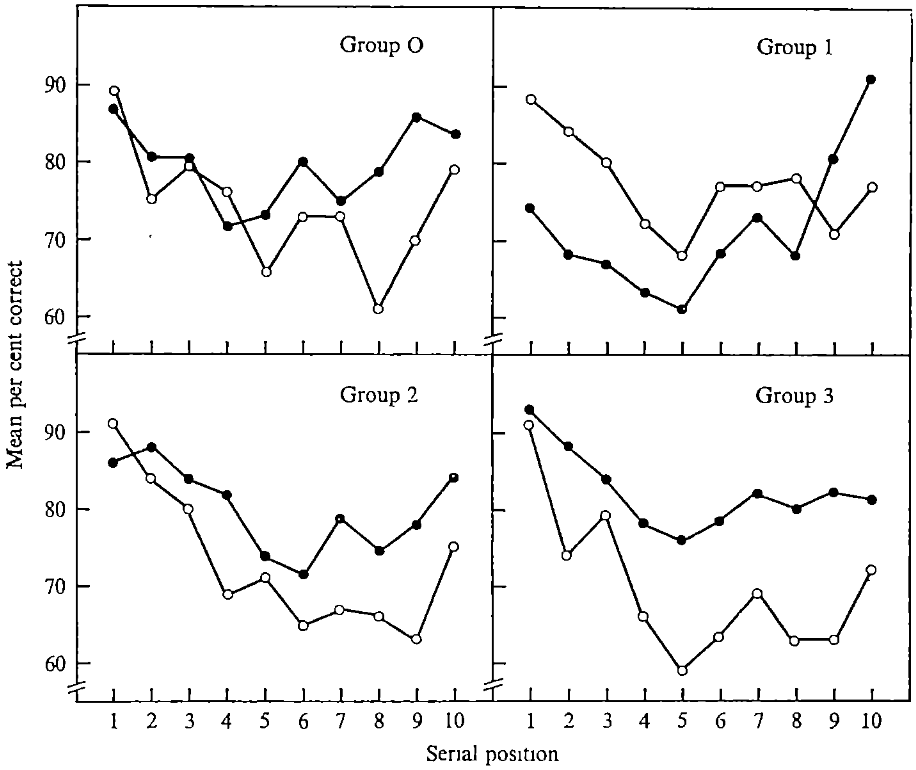


Fig. 2. Serial position curves from Expt. II before (○—○) and after (●—●) instructions to use imagery. Group 0: no instructions. Group 1: separative imagery instructions. Group 2: interactive imagery instructions by pairs of items. Group 3: interactive imagery instructions by groups of items.

540, $P < 0{\cdot}001$), and this also showed a significant interaction with the difference between Group 1 and Groups 2 and 3 ($F = 2{\cdot}67$, d.f. = 9, 540, $P < 0{\cdot}01$). Finally, the interactions between the effect of conditions and that of serial position, and between the effect of conditions, the effect of serial position, and the difference between Group 1 and Groups 2 and 3 were also significant ($F = 4{\cdot}80$, d.f. = 9, 540, $P < 0{\cdot}001$; and $F = 2{\cdot}15$, d.f. = 9, 540, $P < 0{\cdot}05$).

A posteriori tests were carried out on the effect of conditions for each of the four groups. There was a small improvement in Group 0, which approached, but did not reach significance ($F = 3{\cdot}80$, d.f. = 1, 60, $0{\cdot}1 > P > 0{\cdot}05$). There was a significant decrement in Group 1 ($F = 4{\cdot}02$, d.f. = 1, 60, $P < 0{\cdot}05$). There was a significant improvement in Group 2 ($F = 6{\cdot}41$, d.f. = 1, 60, $P < 0{\cdot}02$), and an even larger one in Group 3 ($F = 19{\cdot}83$, d.f. = 1, 60, $P < 0{\cdot}001$). Further tests showed no significant difference between Group 0 and Groups 2 and 3 ($F < 1$), nor any interaction with the effect of conditions ($F = 1{\cdot}59$, d.f. = 1, 60, $P > 0{\cdot}2$), nor any three-way interaction with the effect of conditions and that of serial position ($F < 1$).

Various properties of the subjects' recall performance are summarized in Table 2. The first row shows the mean percentage of occasions when the subject recalled the items in serial position 1 first, and the second row shows the mean percentage of occasions when the subject recalled the item in serial position 10 first. Inspection of

Table 2. *Properties of recall in Expt. II*

Condition	I	II			
Group	0–3	0	1	2	3
Mean % occasions when item at serial position 1 was recalled first	37·1	15·6	14·8	25·8	30·5
Mean % occasions when item at serial position 10 was recalled first	22·3	15·6	35·2	27·3	32·8
Mean number of forward repetitions per list	2·84	4·02	1·85	2·95	3·06
Mean number of backward repetitions per list	0·59	0·43	0·89	0·91	1·15
Mean expected number of repetitions per list	0·64	0·70	0·62	0·70	0·72

Key: Group 0: no instructions. Group 1: separative imagery instructions. Group 2: interactive imagery instructions by pairs of items. Group 3: interactive imagery instructions by groups of items.

Table 2 shows that Groups 2 and 3 tended to recall from both the beginning and the end of the list in Condition II, Group 1 tended to recall from the end of the list, and Group 0 tended not to recall from either the beginning or the end.

The last three rows of Table 2 show a measure of organization adapted from Bousfield & Bousfield (1966). A forward repetition is defined as any pair of items presented in immediate succession and recalled in immediate succession in the order of presentation. A backward repetition is defined as any such pair recalled in the reverse order. Inspection of Table 2 shows that in Condition II the groups given imagery instructions produced more backward repetitions and fewer forward repetitions than the control group; further, the group given separative instructions produced fewer forward repetitions than the groups given interactive instructions.

Finally, one can compute the expected number of repetitions if the items actually recalled were given in a completely random order. This is the same for both forward and backward repetitions, and is shown in the last row of Table 2. Separate two-tailed t tests showed that, in Condition I, the number of forward repetitions was greater than expected, while the number of backward repetitions did not differ significantly from that expected by chance ($t = 12{\cdot}57$, d.f. = 63, $P < 0{\cdot}001$; and $t = 0{\cdot}76$, d.f. = 63, $P > 0{\cdot}4$, respectively). In Condition II, all four groups produced more forward repetitions than would be expected by chance ($t = 7{\cdot}60, 4{\cdot}31, 6{\cdot}05, 6{\cdot}18$, respectively, d.f. = 15, $P < 0{\cdot}001$ in all cases). Group 0 produced a smaller number of backward repetitions than would be expected by chance, Group 1 produced a greater number than would be expected, but Groups 2 and 3 did not differ from chance ($t = 2{\cdot}71$, $P < 0{\cdot}02$; $t = 2{\cdot}87$, $P < 0{\cdot}02$; $t = 0{\cdot}94$, $P > 0{\cdot}2$; $t = 1{\cdot}10$, $P > 0{\cdot}2$, respectively; d.f. = 15 in all cases).

Discussion

Although the subjects who were instructed to use interactive imagery in this experiment produced better performance than those who were instructed to use separative imagery, the former subjects did not produce better performance than those in the control group, who were left to their own resources. While contrary to previous results (Bower & Winzenz, 1970), this is perhaps not surprising. On the one hand, the groups using interactive imagery had only a limited room for improvement; on the other hand, the control subjects were free to employ whatever strategies they chose, and they might well have chanced upon something reasonably effective by the time that they had learned eight lists. In terms of their recall performance, then, the subjects using interactive imagery were very similar to the controls. This is also true of the serial position curves (Fig. 2). Nevertheless, if one considers within-subject comparisons, interactive imagery instructions produced a significant improvement in performance between the first eight lists and the second eight, whereas the control subjects did not show a significant improvement. The increase in performance in Groups 2 and 3 appears to apply over all serial positions, and so in terms of the PM/SM distinction interactive imagery relates to SM, rather than to PM.

The subjects who were instructed to use separative imagery produced quite different results from the other three groups. As a result of their instructions, they produced a significant decrement in performance, and an interaction with the effect of serial position. This interaction took the form of an increased recency effect and a reduced primacy effect. This was precisely the outcome which was predicted earlier, on the basis of the idea that subjects of high spatial ability use separative imagery in learning. It also constitutes yet another result relating the various operational procedures for investigating imagery.

In the discussion of Expt. I, it was suggested that Craik's (1969) results on the effect of instructions to recall in a particular order might be invoked to explain the shape of the serial position curve produced by subjects of high spatial ability. Similarly, it might be proposed that control subjects and those given interactive imagery instructions tend to recall from the beginning of a list (producing a large primacy effect and a small recency effect), whereas subjects given separative imagery instructions tend to recall from the end of the list (producing a small primacy effect and a large recency effect). Unfortunately, the results in Table 2 concerning the serial position from which recall commences fail to support this analogy. The subjects who were given separative imagery instructions did tend to recall from the end of a list, but those who were given interactive imagery instructions tended to recall from both the beginning and the end, while the control subjects tended not to recall from either the beginning or the end. Moreover, while it was true that the subjects who were given separative imagery instructions tended not to produce forward repetitions, the subjects who were given interactive imagery instructions were also less likely to do so than the controls.

The effects of instructions upon the shape of the serial position curve can perhaps be best handled by supposing that both separative imagery instructions and instructions to commence recall from the end of a list disrupt the subject's organization of the list; whereas interactive imagery instructions and instructions to commence recall from the beginning of a list promote his organization of the list. The present results support this idea if one can suppose that the total number of repetitions (forward and backward) is a measure of the extent to which the subject has been able to organize the list. Inspection of Table 2 shows that the subjects who were given separative imagery instructions were clearly poorer than the other three groups on this measure.

Although the control subjects produced similar results to those who were given interactive imagery instructions in terms of their overall performance, there are several differences in terms of the variables considered in Table 2. In particular, the control subjects had a much greater tendency to produce forward repetitions than any of the other groups, that is, they tended to recall in the original order of presentation. One possible explanation for this is that it reflects the use of verbal mediators, perhaps the production of sentences containing the words to be remembered. The linear constraints of linguistic organization would produce a large number of forward repetitions, and an even smaller number of backward repetitions than would be expected on the basis of a random ordering of the responses. This was found in all of the groups before they were given differential instructions, and in the control group on all 16 lists. One result which is consistent with this explanation is that explicit instructions to construct sentences produce similar performance to instructions to use imagery, but they lead to better organization (Bower & Winzenz, 1970).

CONCLUSIONS

Together with the results of previous studies (Richardson, 1974, 1975*a*, *b*), the experiments described here enable one to make certain statements concerning the procedures for studying imagery. If subjects are given standard free recall instructions, the imageability of the stimulus material will affect their performance on items of low concreteness, but not on items of high concreteness. The positive effect of imageability on items of low concreteness will be greater in the case of those subjects who perform well on tests of spatial ability. If subjects are given instructions to use imagery in learning, the imageability of the stimulus material will affect their performance on items of both high and low concreteness. It may be concluded that subjects given standard instructions only use imagery to remember items of low concreteness, but that subjects given instructions to use imagery to remember all items will do so.

The positive effect of imageability upon the recall of items of low concreteness has its locus in secondary memory, not in primary memory. The positive effect of instructions to use interactive imagery also applies to the SM component (see also Smith *et al.* 1971). However, the effect of imagery ability and the effect of instructions to use separative imagery do not fit readily into the PM/SM dichotomy. Subjects of high imagery ability and subjects who are given separative imagery instructions produce a larger recency effect and a smaller primacy effect. It may be concluded that subjects of high imagery ability use imagery in learning, but that they do not use interactive imagery. This leads to poorer organization, and no improvement in overall performance.

The original exponents of the PM/SM dichotomy attempted to find manipulations which would affect one of the components at a time (e.g. Glanzer & Cunitz, 1966). This approach does not enable one to give a sensible account of results showing effects on both components simultaneously. However, if such effects can be related to a disruption of organization, one may offer explanations in terms of the alternative conceptions of free recall performance which were described previously. If a subject's organization of a list is disrupted, this might well lead to a greater use of a short-term store as a register, on the account offered by Rundus (1971), or to a greater use of an ordinal retrieval strategy, on the account offered by Baddeley & Hitch (1974). The present results do not permit a clear decision between these two approaches, and it would probably be more profitable to try to distinguish between them in other experimental situations. However, insofar as the repetition measure used in Expt. II is not only a measure of long-term organization, but also of the extent to which the subjects employed an ordinal retrieval strategy, the results are in the opposite direction to that which would be predicted on the latter interpretation, in favour of the account given by Rundus. This account was based upon the dual-storage model proposed by Atkinson & Shiffrin (1968), which made an explicit distinction between two functions of the short-term store: the maintenance of information in a temporary register, and the transfer of information to long-term memory. This has been expressed more recently in terms of the 'maintenance' aspect of rehearsal and the 'elaborative' or 'constructive' aspect (Craik & Watkins, 1973; Woodward, Bjork, & Jongeward, 1973).

While a distinction between primary and secondary memory is able to give a meaningful and interesting account of some findings, in terms of specific effects upon one or other component of free recall performance, it is not adequate to handle other results, such as those presented here. It would seem to be profitable to extend the dichotomy in the direction pursued by Atkinson & Shiffrin, so as to give due regard to the variety of strategies which may be adopted by subjects in learning experiments.

This research was supported by the Science Research Council and the Medical Research Council. The experiments were carried out while the author was at the Laboratory of Experimental Psychology, University of Sussex. I am grateful for the helpful advice and criticisms of Alan Baddeley, Allan Paivio, and Tim Shallice.

REFERENCES

ATKINSON, R. C. & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (eds), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. II. New York: Academic Press.

BADDELEY, A. D. & HITCH, G. (1974). Working memory. In G. H. Bower (ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. VI. New York: Academic Press.

BETTS, G. H. (1909). *The Distribution and Functions of Mental Imagery*. New York: Teachers College.

BOUSFIELD, A. K. & BOUSFIELD, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychol. Rep.* **19**, 935–942.

BOWER, G. H. (1970). Imagery as a relational organizer in associative learning. *J. verb. Learn. verb. Behav.* **9**, 529–533.

BOWER, G. H. (1972). Mental imagery and associative learning. In L. W. Gregg (ed.), *Cognition in Learning and Memory*. New York: Wiley.

BOWER, G. H. & WINZENZ, D. (1970). Comparison of associative learning strategies. *Psychon. Sci.* **20**, 119–120.

CLARK, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *J. verb. Learn. verb. Behav.* **12**, 335–359.

CRAIK, F. I. M. (1969). Modality effects in short-term storage. *J. verb. Learn. verb. Behav.* **8**, 658–664.

CRAIK, F. I. M. & WATKINS, M. J. (1973). The role of rehearsal in short-term memory. *J. verb. Learn. verb. Behav.* **12**, 599–607.

DI VESTA, F. J., INGERSOLL, G. & SUNSHINE, P. (1971). A factor analysis of imagery tests. *J. verb. Learn. verb. Behav.* **10**, 471–479.

GLANZER, M. & CUNITZ, A. R. (1966). Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.* **5**, 351–360.

HOLLENBERG, C. K. (1970). Functions of visual imagery in the learning and concept formation of children. *Child. Dev.* **41**, 1003–1015.

MARKS, D. F. (1972). Individual differences in the vividness of visual imagery and their effect on function. In P. W. Sheehan (ed.). *The Function and Nature of Imagery.* New York: Academic Press.

PAIVIO, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

PAIVIO, A., YUILLE, J. C. & MADIGAN, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. exp. Psychol.* (Mongr. Suppl.) **76**, 1–25.

RICHARDSON, J. T. E. (1974). Imagery and free recall. *J. verb. Learn. verb. Behav.* **13**, 709–713.

RICHARDSON, J. T. E. (1975*a*). Imagery, concreteness, and lexical complexity. *Q. Jl exp. Psychol.* **27**, 211–223.

RICHARDSON, J. T. E. (1975*b*). Concreteness and imageability. *Q. Jl exp. Psychol.* **27**, 235–249.

RUNDUS, D. (1971). Analysis of rehearsal processes in free recall. *J. exp. Psychol.* **89**, 63–77.

SHEEHAN, P. W. (1966). Functional similarity of imaging to perceiving: Individual differences in vividness of imagery. *Percept. mot. Skills* **23**, 1011–1033.

SHEEHAN, P. W. (1967*a*). A shortened form of Betts' Questionnaire Upon Mental Imagery. *J. clin. Psychol.* **23**, 386–389.

SHEEHAN, P. W. (1967*b*). Reliability of a short test of imagery. *Percept. mot. Skills* **25**, 744.

SHEEHAN, P. W. & NEISSER, U. (1969). Some variables affecting the vividness of imagery in recall. *Br. J. Psychol.* **60**, 71–80.

SMITH, E. E., BARRESI, J. & GROSS, A. E. (1971). Imaginal versus verbal coding and the primary–secondary memory distinction. *J. verb. Learn. verb. Behav.* **10**, 597–603.

STEWART, J. C. (1965). An experimental investigation of imagery. Unpublished doctoral dissertation, University of Toronto.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of 30,000 Words.* New York: Teachers College.

THURSTONE, L. L. & JEFFREY, T. E. (1956). *Flags: A Test of Space Thinking.* Chicago: Education Industry Service.

TULVING, E. (1968). Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (eds), *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, N.J.: Prentice-Hall.

WAUGH, N. C. & NORMAN, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

WOODWARD, A. E., JR., BJORK, R. A. & JONGEWARD, R. H., JR., (1973). Recall and recognition as a function of primary rehearsal. *J. verb. Learn. verb. Behav.* **12**, 608–617.

(*Manuscript received* 17 *April* 1975; *revised manuscipt received* 30 *July* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 501–506
Printed in Great Britain

THE RELATION OF REPORTED IMAGERY TO COGNITIVE PERFORMANCE

By A. J. DURNDELL
Department of Social Sciences, Glasgow College of Technology
and N. E. WETHERICK
Department of Psychology, University of Aberdeen

Six groups of subjects with known scores on reported controllability and vividness of imagery were tested on five cognitive tasks. A high 'controllability of visual imagery' group performed two divergent thinking tasks better than a low 'controllability' group. Self-rated use of imagery was positively associated with performance on the divergent thinking tasks and negatively associated with performance on a Gestalt-type problem task. No relationships were found between reported imagery and performance on two concept tasks, one of which was purely visual.

The recent revival in studies of imagery has tended to concentrate on the role of imagery in memory and paired associate learning, following the work of Paivio (e.g. Paivio, 1971). There have been some notable exceptions such as Piaget & Inhelder (1971), who tied imagery into their general cognitive developmental framework, but on the whole there have been few studies of the relationship between imagery and other thought processes. Thus it is not known whether imagery has any useful part to play in cognitive performance.

Paivio himself has tended recently to move away from the learning paradigm: for example, Paivio (1973) discusses a dual processing model of human thought with imagery playing an important part on the non-verbal side of thinking rather than the verbal. There has also been a tendency to view imagery as more involved with the less logical side of human thinking. An example is provided by McKellar (1968), who compares 'A-thinking' which he sees as less logical, possibly more divergent and less verbal in character, with 'R-thinking' which he sees as logical, convergent and verbal. Discussion of this kind has been somewhat speculative since few studies have tried to evaluate the suggested relationships.

The interchangeability of the terms 'divergent thinking' and 'imagination' seems to imply that there is some relationship between imagery and divergent thought, and there is anecdotal evidence of the use of imagery in divergent thought which at least suggests that the two processes can be combined in certain circumstances. One well-known example is provided by Enid Blyton, who wrote her books from an ongoing stream of images, and another by William Blake who sketched or painted his 'visions'. Related evidence concerning the role of imagery produced by hallucinogenic drugs in divergent thought has hardly progressed beyond the anecdotal stage (cf. Tart, 1969). On a simpler level, Hargreaves & Bolton (1972) found that a measure of mnemonic paired associate learning (using visual imagery) related quite highly to both divergent thinking tasks (e.g. uses of objects) and convergent (IQ) tasks. The correlations were highest with the divergent tasks whilst the factor loadings were highest on an 'IQ' factor. On the basis of this, it could be predicted that performance on divergent thinking tasks will be positively related to vividness and controllability of imagery as subject variables.

Richardson (1969) discussed the relationship between imagery and problem solving in the sense used by the Gestalt psychologists. He argued that uncontrolled vivid imagery could be a hindrance to problem solving by getting in the way and disrupting a profitable line of thought, but that imagery, particularly controlled vivid imagery, might also aid problem solving by providing an alternate mode of thinking to verbal thought. In the latter case it could help by breaking an ineffective set and causing the elements of the problem to be 'restructured' or 'reorganized', to use Gestalt terms. The present study aimed to see whether either of these two statements could be supported.

The relationship between image and concept is very complex. Both Arnheim (1970) and Horowitz (1970) discuss the issue at length, and appear to break down the distinction between concrete and abstract in this context. They are also unhappy about the association of imagery with concreteness and concepts with abstractness. There is little research in this area, although Kuhlman (1960) and Hollenberg (1970), using four spatial tests to measure imagery in young children, found that the low 'imagery' groups were significantly better than the high 'imagery' groups at learning verbal concepts, the difference decreasing with age (6–10 years). One is almost left with Galton's (1883) contention that imagery is a hindrance in dealing with abstract concepts, and that as professors deal with abstract concepts, this explains why professors have little imagery! It could be predicted, however, that vivid and controllable imagery would be useful for identifying a visual concept and a hindrance in identifying a verbal concept.

Richardson (1969) contends that two self-report questionnaires tap the most important dimensions of imagery, these being Sheehan's (1967) modification of Bett's Vividness of Imagery (VI) questionnaire and Gordon's (1949) Controllability of Visual Imagery (CI) questionnaire. Consequently these questionnaires were used to assess subject differences on these two dimensions.

METHOD

Subjects

60 subjects were chosen from 279 who had filled in the CI and VI questionnaires. The distribution of the CI scores was divided at the mean and the distribution of the VI scores divided into three equal parts, the resulting two-dimensional plot yielding six cells. The ten subjects for each of the cells formed were chosen on the basis of being farthest from the boundary scores for the cell. Response rate was approximately 80 per cent and replacements were obtained where necessary utilizing the same criterion. Mean score (no. of items controlled) for the high CI group was 11·68, and for the low CI group 6·35. Mean VI scores (average rating of vividness of suggested images) for the high, medium and low VI groups were 1·81, 2·71 and 3·98 respectively (low VI score indicated high vividness of imagery).

Tasks

Imagery questionnaires: Both questionnaires were taken directly from Richardson (1969) and were stapled together into one booklet.

Divergent thinking tasks: sheets of unlined paper with the word SHOE or with ○ ○ (two circles) printed on them.

Problem-solving task: a variety of objects, e.g. drawing pins, pencils, comb, magazine, including paper clips and a piece of white card, scattered on a table. There was a nail stuck in the wall of the experimental room at head height.

Concept tasks: Twenty sets of printed Chinese characters. Each set comprised four examples of Chinese characters which had a common root character plus four further examples in an array only one of which had the same root as the first four examples.

Twenty sets of groups of English letters printed on cards. Each set comprised three groups of five letters, the groups having two letters in common, plus a further four groups in an array only one of which contained the two letters common to the first three groups.

These two types of task were chosen in order to provide one (the Chinese characters) which was purely visual with no possibility of the subject being able to name the common element, and another (the English letters) which was partially visual and partially verbal and in which the common elements could be named allowing the use of either visual or verbal strategies.

Procedure

The two imagery questionnaires were completed under group conditions. Verbal and written instructions placed particular stress on the difference between knowing something abstractly and being able to image something. The instructions were designed to avoid giving experience of imagery either a positive or a negative evaluation.

Each experimental subject was tested individually and the tasks were presented in a fixed order. Within each of the six cells subjects were randomly allocated to an 'imagery instructions' or a 'no imagery instructions' condition, the former being instructed to use imagery whenever possible. A sheet of paper with SHOE printed on it was then produced, and the subject was given 5 min to write down as many uses for a shoe as he could think of. This was then repeated with the two circles (○ ○), with instructions to draw as many objects incorporating the two circles as possible, labelling each drawing after completing it. The suggestion that the circles might represent wheels was given to all subjects. Two marks were produced by an independent marker for the divergent thinking tasks. The first was a total mark (fluency) not allowing absolute repetitions, the second (unusual uses) allowed only one use for uses of a shoe directly concerned with wearing on the feet, and only one use for two circles representing wheels on transport. The marker had the power to give credit for unusual suggestions within these categories.

The subjects were next instructed to hang the white card from the nail in the wall, utilizing anything present on the table with the constraint that no hole was to be made in the card. The easiest way to do this was to use the paper clip, which involved perceiving an unusual use for it. In Gestalt terms this can be seen as 'restructuring' the problem. Time to solution was recorded.

The subjects were then presented with the 20 Chinese character trials; on each trial, the examples with a common element were shown sequentially, followed by the array of four further examples from which the subject had to choose the item including the common element. The process was repeated with the 20 sets of English letters. Subjects were awarded one mark for each correct answer, giving two total marks out of 20. Finally, each subject was asked to rate, on a seven-point scale, his use of imagery in each of the five tasks, after being reminded what the tasks were.

RESULTS

Analyses of variance, the main effects being CI by VI by instructions to use imagery, were calculated on the performance scores for both divergent thought tasks. The major finding was a significant superiority of the high CI group over the low CI group on three of the four divergent thinking measures, and its non-significant superiority on the fourth (see Table 1). No other effects or interactions were significant. All the divergent thinking scores behaved in an approximately similar fashion, supporting Hargreaves & Bolton's (1972) conclusion that it is redundant to use more than a few measures of this type of thinking. The ratings given by subjects of their use of imagery whilst performing the divergent thinking tasks, correlated

Table 1. *Performance of the CI groups on the divergent thinking tasks*

	Uses of shoe				Uses of ○ ○			
	Fluency		Unusual uses		Fluency		Unusual uses	
Group	Score	Significance	Score	Significance	Score	Significance	Score	Significance
High CI ($n = 30$)	8·43	n.s.	6·60	$P < 0{\cdot}01$	7·97	$P < 0{\cdot}001$	7·43	$P < 0{\cdot}01$
Low CI ($n = 30$)	6·90		4·47		5·93		5·53	

Table 2. *Correlations between performance scores and use of imagery ratings* ($n = 60$)

	Divergent thinking task					Concept tasks	
	Uses of shoe		Uses of ○ ○		Problem-solving task	Chinese characters	English letters
	(I)	(II)	(I)	(II)			
Correlation	0·5848	0·3489	0·3595	0·2900	−0·3123	0·1296	0·0446
Significance	< 0·001	< 0·01	< 0·01	< 0·05	< 0·02	n.s.	n.s.

(I) = fluency; (II) = unusual uses.

significantly and positively with performance on the divergent thinking tasks for all four measures (see Table 2).

Analysis of performance scores on the problem task, in terms of the numbers of subjects in each group who took above 25 sec to solve the problem (i.e. whether or not the problem was solved relatively quickly) produced no significant results. However, time taken to solution correlated significantly and negatively with rated use of imagery during attempted solution, i.e. subjects who took longer tended to use more imagery. This does not say how useful the imagery was, as the process leading to insightful solution could actually be hindered by use of imagery and give this result.

Analyses of variance, the main effects being CI by VI by instructions to use imagery, were calculated on the performance scores on the concept tasks. No significant results were obtained, nor did the ratings of use of imagery during performance on the concept tasks correlate significantly with the performance itself.

Discussion

The clearest finding was related to the prediction that imagery as a subject variable would be positively related to divergent thought production. Self-rated vividness of imagery over all sense modalities (VI) was not found to be related but both self-rated controllability of visual imagery (CI) and self-rated use of imagery during performance were. This indicates that subjects felt that they tended to use imagery when successfully thinking of uses for objects, and that those who were able to control their visual imagery were at an advantage. The visual dimension would seem to be relevant (only one-seventh of the vividness questionnaire's items were concerned with vision, which may account for the failure to find a relationship). McKellar (1965) and others have found the visual to be the most commonly occurring imagery modality, and this could mean that it is the most useful form of imagery to maintain some degree of control over. These findings largely agree with Hargreaves & Bolton's (1972) finding of an association between divergent thinking and facility at learning paired associates of high imagery value, and with the anecdotal evidence that imagery can play a part in the divergent process. If the most productive approach is one which maximizes the switching of styles, for instance between a verbal style and an imagery style, then visual imagery which is guided or controlled can be expected to be of more use than autonomous imagery which is not under conscious control. This does seem to indicate that imagery has a useful role to play in some thought processes and is not a useless by-product and that both sides in Paivio's dual processing model can be used in divergent production as they have already been shown to be used in memory. At the same time the connexion between imagery and divergent thought makes the inclusion of both on the A-thinking side of McKellar's dichotomy of thought processes seem more reasonable.

Neither of Richardson's (1969) two predictions (that uncontrolled vivid imagery could be a hindrance in problem solving and that controlled vivid imagery could be helpful) were borne out. No relationships at all were found. The finding that the longer subjects took to solve the Gestalt problem (i.e. the worse their performance), the more imagery they reported using, implies that if the problem causes difficulty the switching of styles approach is used. Whether the actual use of imagery in this context aids or hinders solution cannot be ascertained from this study, only that it is reported to be used more when performance is worse.

The predictions that VI and CI would be useful on a completely visual concept task (the Chinese characters), and might be a hindrance on a more verbal concept task (the English letters), were not borne out. No relationship at all was found between performance on these tasks and either VI and CI groups or rated use of imagery. This indicates the great care that must be taken in equating high reported conscious imagery with ability on a visual task (e.g. the Chinese characters), a point that may invalidate the use of spatial intelligence test scores as criteria of imagery (cf. Hollenberg, 1970). The most likely explanation for the lack of findings would seem to be that in both tasks identification of the common element was a relatively immediate perceptual process which did not require the formation of conscious imagery. Many subjects thought that they were guessing when in fact they were achieving 75 per cent accuracy. Identification of a common visual element involves visual abstraction but not necessarily any conscious visual representation, so the role of imagery in such a concept task could be very minor. Kuhlman (1960) and Hollenberg (1970) used materials that involved verbal concepts related to concrete examples, which could possibly offer a role to conscious imagery, but use of imagery groups based on spatial intelligence test performance would not necessarily give the same results as using groups chosen on the basis of reported conscious imagery.

Several other points can be made about the procedure and the assessments of imagery. It might be argued that fixed order of presentation of the tasks was a source of bias in the experiment. However, the object was to look separately at the subjects' performance on each of the five tasks to see whether imagery status affected their performance, not to compare performance on the different tasks. Furthermore, the tasks were all short and different from each other and far from becoming fatigued, subjects found them novel and interesting.

As regards the assessment of imagery, the results obtained from the ratings of use of imagery during performance (see Table 2) indicate that it is not the case (as has been suggested) that subjects rate their use of imagery highly only when they perform well on a task. For instance use of imagery was positively related to successful divergent production and negatively related to successful problem solving. This gives the ratings some face validity, as it appears that subjects are able to discriminate between different subjective levels of use of imagery on different tasks. The VI questionnaire failed throughout to provide subject groups who differed in performance, and it would seem from these results that reported vividness of imagery measured over all modalities may not be a worthwhile dimension to investigate in this context. Controllability of visual imagery and reported use of imagery seem more directly relevant, although it is possible that a *visual* vividness questionnaire might be useful. Finally, instructions to use imagery proved ineffectual in altering performance on the tasks, even where imagery was found to be significantly related to performance. This implies that imagery is used and people are aware of using it without prompting.

The research reported in this paper was carried out whilst A. J. Durndell held a Departmental Studentship at the Psychology Department, Aberdeen University.

REFERENCES

ARNHEIM, R. (1970). *Visual Thinking*. London: Faber & Faber.

GALTON, F. (1883). *Enquiries into Human Faculty*. London: Macmillan.

GORDON, R. (1949). An investigation into some of the factors that favour the formation of stereotyped images. *Br. J. Psychol.* **39**, 156–167.

HARGREAVES, D. J. & BOLTON, N. (1972). Selecting creativity tests for use in research. *Br. J. Psychol.* **63**, 451–462.

HOLLENBERG, C. K. (1970). The functions of visual imagery in the learning and concept formation of children. *Child Dev.* **41**, 1003–1014

HOROWITZ, M. J. (1970). *Image Formation and Cognition*. New York: Appleton Century Crofts.

KUHLMAN, C. (1960). Visual imagery in children. Unpublished doctoral dissertation, Harvard University. Discussed in J. S. Bruner *et al.* (eds), *Studies in Cognitive Growth* (1966). New York: Wiley.

MCKELLAR, P. (1965). The investigation of mental images. In S. A. Barnett & A. McLaren (eds), *Penguin Science Survey, 1965 B*. Harmondsworth: Penguin.

MCKELLAR, P. (1968). *Experience and Behaviour*. Harmondsworth: Penguin.

PAIVIO, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.

PAIVIO, A. (1973). Psychophysiological correlates of imagery. In F. J. McGuigan & R. Schoonover (eds), *The Psychophysiology of Thinking*. New York: Academic Press.

PIAGET, J. & INHELDER, B. (1971). *Mental Imagery in the Child*. London: Routledge & Kegan Paul.

RICHARDSON, A. (1969). *Mental Imagery*. London: Routledge & Kegan Paul.

SHEEHAN, P. W. (1967). A shortened form of Bett's questionnaire upon mental imagery. *J. clin. Psychol.* **23**, 386–389.

TART, C. T. (ed.) (1969). *Altered States of Consciousness*. New York: Wiley.

(*Manuscript received* 12 *February* 1975; *revised manuscript received* 28 *August* 1975)

Br. J. Psychol. (1976), 67, 4, *pp*. 507–514
Printed in Great Britain

VISUAL BEHAVIOUR IN SOCIAL INTERACTION: THE VALIDITY OF EYE-CONTACT ASSESSMENTS UNDER DIFFERENT CONDITIONS OF OBSERVATION

BY BERNARD RIMÉ
University of Louvain

AND LEX MCCUSKER
Dartmouth College

On each session of the experiment, each of five subjects was assigned to act as either gaze-sender (GS), gaze-receiver (GR), direct-sight observer (DSO), TV observer (TVO) or one-way mirror observer (OWO). The GS had to look into the GR's eyes each time he received pre-recorded signal through earphones. The GR and the three observers recorded GS's eye-gaze behaviour. Two conditions of distance between GR and GS were included. Results showed that (1) the GR, DSO and TVO presented an equivalent accuracy in their EC recordings as compared with the programmed signals, while the OWO's accuracy appeared significantly lesser; (2) as a concealed observer, OWO was also less efficient than TVO since the agreement of the former with GR's records was significantly lower than that of the DSO, which was not the case for the latter; (3) with increment of distance, OWOs recorded longer EC periods, a tendency which was not found significant among the other observers.

Studies on visual behaviour in social interaction were virtually non-existent before the sixties, but a growing interest in that field has arisen during the past few years. Eye-contact between partners and gaze directed toward partner's eyes have become widely used dependent variables in experimental research on dyadic interaction. As pointed out by Efran (1968), investigators have been interested in visual interaction both because of the central role visual cues play in non-verbal communication and because of the potential use of eye-contact patterns as an unobtrusive measure of other variables. However, it presumes that eye-contact and visual orientation toward partner's eyes can be correctly assessed to a high degree.

In most studies where eye-contact served as a dependent variable (e.g. Exline, 1963; Strongman & Champness, 1968; Goldberg, Kiesler & Collins, 1969; Rutter & Stephenson, 1972), the accuracy of gaze assessments was tested by the use of independent observers and the computation of an interobservers' agreement index. However even when such agreement appears to be high, which is not always the case (e.g. Mehrabian & Williams, 1969; Vine, 1971; Stephenson, Rutter & Dore, 1972), the problem of judgement validity, that is, the correspondence between signals which have been sent and signals which have been received, remains unresolved. This constitutes an experimental problem in itself and implies the use of a paradigm where the sent signals are controlled. The looking behaviour of the sender being under control of the experimenter, it becomes possible to estimate the accuracy of gaze assessments by the receiver himself and by observers not involved in the dyadic interaction, as well as the correspondence between receiver's and observer's recordings. Studies of that kind (e.g. Gibson & Pick, 1963; Cline, 1967; von Cranach, Hückstedt, Schmid & Vogel, 1968; Anstis, Mayhew & Morley, 1969; Krüger & Hückstedt, 1969; Ellring & von Cranach, 1972) have been extensively reviewed by von Cranach &

Ellring (1973). These authors concluded that 'a number of factors could be identified such as gaze movement, head direction, head movement, distance, duration and personality variables that impair the judgement of the receiver'. As far as the observers are concerned, von Cranach & Ellring noted that, generally, they discriminated much worse than did the receivers, a number of additional factors influencing their observations, so that 'the validity of the observer's judgement depends predominantly on the experimenter's choice of specific observational circumstances'.

A particular aspect of the choice of the observational circumstances in an experiment using visual behaviour as a dependent variable is the location and concealment of the observer. His presence in the experimental room itself being rarely suitable because of the highly probable obtrusive effects it would have in experiments on verbal and non-verbal interaction, two kinds of media are commonly used: the one-way mirror and the closed circuit television. The present study is an attempt to compare the accuracy of eye-directed gaze recordings simultaneously performed by the receiver himself, by a direct-sight observer, by an observer behind a one-way screen and by an observer having a TV screen view of the sender's face. In order to get the criterion for establishing the validity of each of the recordings, the visual behaviour of the sender was programmed throughout the experiment. Because of the importance of distance as a factor in the recognition of gaze direction (Stephenson & Rutter, 1970; Argyle, 1970; Stephenson *et al.* 1972; Knight, Langmeyer & Lundgren, 1973; von Cranach & Ellring, 1973), two levels of distance between the sender and the receiver and observers' places were included in the design, resulting in a 2 × 4 between-subjects factorial design.

METHOD

Subjects

Through 18 experimental sessions, a total number of 90 subjects were used. They were male and female volunteer undergraduate students enrolled in introductory psychology courses. On each session, five subjects were randomly assigned to one of five different roles: one of them had to act as gaze-sender (GS), another as gaze-receiver (GR) and the three remaining as direct-sight observer (DSO), TV observer (TVO) and one-way mirror observer (OWO).

Experimental setting

The experiment was conducted within a 12 × 24 ft experimental room provided with five external adjacent cubicles. A one-way mirror separated one of these cubicles from the experimental room. The GS and the GR sat facing each other, the GR having his back close to the mirror. The distance between them was 6 ft in the smaller condition of distance and 9 ft in the larger. The DSO sat next to the GR, about 1 ft to his left. The OWO was seated behind the one-way mirror so that he could observe the GS's face between the backs of the GR and the DSO. A distance of 2 ft, due to wall thickness, separated the OWO from the GR and DSO. A camera provided with a 25–100 mm $f/1{\cdot}8$ zoom lens was mounted on the wall of the experimental room, just above the GR's head and focused on the GS's face (Fig. 1).

Through earphones, the GS received programmed signals for the control of his gaze direction. He was instructed to direct his eyes right into the GR's eyes each time a tone was heard in the earphone. During silent periods he had to look 'anywhere on the wall in front of him, except in his partner's eyes'. Each time an event occurred, i.e. eye-directed gaze from the GS toward the GR, the GR himself and the three observers depressed a button to activate the corresponding channel of an event-recorder. One of the cubicles served as the control room in which were located a cassette tape-recorder for sending the gaze direction signals, a five-channel event-recorder and a TV monitor to control the image taken by the camera. Through four channels of the event-recorder, at a chart speed of 3 mm/sec, a continuous recording of the four subjects' observations about GS's gaze direction was made. By means of an electronic cell connected at the

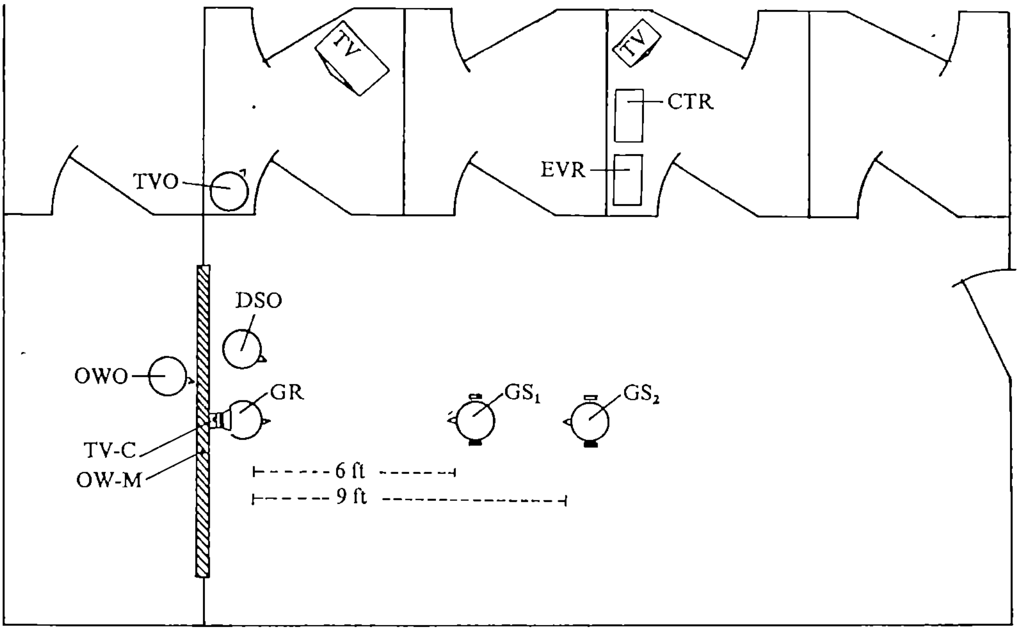


Fig. 1. Laboratory arrangement showing the places of the gaze-sender in the shorter (GS$_1$) and in the longer distance condition (GS$_2$), the places of the gaze-receiver (GR), direct-sight observer (DSO), one-way mirror observer (OWO) and TV observer (TVO). Technical settings comprised a one-way mirror (OW-M), a zoom lens TV camera (TV-C), two TV monitors (TV), a five-channel event-recorder (EVR) and a cassette tape-recorder (CTR).

tape-recorder's output, the pen of the fifth channel was activated by the tone of the eye-contact program given to the GS. With such a system, the onset and duration of EC signals could accurately be paralleled with the onset and duration of GS's eye-contacts as judged by the GR and observers.

Eye-contact programs

Tone signals of constant intensity alternating with silent periods were recorded on cassette tape-recorder to constitute EC programs. Ten programs, each of a length of 3 min, were created. All had the same structure for signal frequency and duration (1 × 15 sec; 1 × 10 sec; 1 × 7 sec; 3 × 5 sec; 1 × 4 sec; 3 × 3 sec; i.e. a total of 60 sec of signals, or one-third of the program duration), but they differed in the signals distribution among silent periods. The proportion of one-third of signals for the 3 min period was inspired by Kendon's (1967) findings on the mean proportion of time spent in mutual gaze during two-person conversations. In each program, three short signals signified the beginning and end of the 3 min session. One program was used each session and the ten programs were rotated through the experiment according to a predetermined random sequence.

Procedure

Upon entering the room, the five subjects were introduced to each other by the experimenter and were told: 'In this experiment, we are interested in the observation of visual behaviour. We are trying to determine the degree of accuracy with which an individual can say that another is, or is not, looking into his own eyes. Moreover, we would like to know how accurately external observers using different observational systems are able to record such eye-contact behaviour. Therefore, you will be randomly assigned to the role of GS, or GR, or observer.'

After drawing lots for the roles, the GS was given his earphones, shown his place and instructed that each time he heard a signal and during its whole length, he would have to look right into the GR's eyes. When no signal was heard, he would direct his gaze, without moving his head, anywhere onto the limits of the wall in front of him, except into the GR's eyes. The GR and the observers were instructed to detect as accurately as possible these periods of EC between GS and GR and to record it by depressing a switch, GR and DSO were instructed to conceal their finger movements to avoid any mutual influence. The several observers were then shown their places. The camera was focused and zoomed on the GS's face and after a short blank trial, a signal was given and the program was sent to the GS.

Scoring system

Using the millimetred scale of the recording chart, it was possible to evaluate by steps of 0·33 sec the agreement between the records of each observer and those of the corresponding programs. Obviously, a simple step-by-step agreement index, i.e. the percentage of overlap of records in time, as often used in EC studies, was inadequate in the present case. Indeed, if an observer had depressed his switch during the whole session, such an index would credit him with a 33 per cent agreement, while another subject who never depressed the switch would get a 66 per cent agreement, because of the program structure. It was therefore indispensable to use an index taking the disagreements into account.

Very good indices of that kind have been developed for assessing interobserver agreement (Light, 1971). Cohen's (1960) K-coefficient appeared appropriate for the present purpose. This coefficient works on the $C \times C$ table resulting from the comparison of two different assignments of items among C categories. It compares the observed entries on the main diagonal with the expected entries on this diagonal. K takes the value of zero when observed agreement equals expected agreement. It takes the value 1·0 when all responses fall on the main diagonal (total agreement). It becomes negative when the observed responses have less than 'chance' agreement.

Above these validity coefficients a classical step-by-step agreement index was established for the record of each of the three observers as compared with those of the GR. These agreement indexes were used to assess how accurate the different observers were in trying to put themselves in the place of the gaze-receiver. Measures of the frequency of recorded EC and of the total length of recorded EC were also taken.

RÉSULTS

Because of a misunderstanding of the instructions, four subjects, two of them acting as TVOs, one as OWO and one as DSO, failed to depress the switch during the experiment. These defections were considered to be randomly distributed across conditions and these subjects were dropped from the data.

The analysis of variance for the validity coefficients (Table 1) revealed no main effect for distance nor significant interaction between distance and kind of observational setting. However, a strong main effect for observers appeared ($F = 10·47$, d.f. = 3, 60, $P < 0·001$). Multiple comparisons of the means performed by Duncan's test showed this effect to be due to a significantly lower K-coefficient among OWOs for both conditions of distance. Among GRs, DSOs and TVOs, K-coefficients were revealed to be homogeneous. Moreover, one may emphasize the fact that the K-coefficients observed in these groups are high. As a point of comparison, let us say for instance that an observer who correctly identified 85 per cent of the programmed EC periods and 85 per cent of the programmed silent periods would have obtained a K-coefficient of 0·67. Records of such an accuracy would have been very satisfying in the present study, taking into account the facts that the GSs only moved their eyes and never their heads, that the observers used were absolutely untrained, that their records were analysed by steps of one-third of a second and that no corrections were introduced in this analysis, for GSs' reaction time to the sonorous signals or for observers' reaction time to GSs' eye movements.

When the GRs' recordings are taken as criterion (Table 2) differences appeared between the step-by-step agreement indices computed for each of the three different

Table 1. *Mean coefficients of validity (K-coefficients) for EC recordings*

	GR	DSO	TVO	OWO
Shorter distance	0·765	0·717	0·600	0·427
Longer distance	0·699	0·691	0·630	0·407

Table 2. *Mean step-by-step agreement with gaze-receiver*

	DSO	TVO	OWO
Shorter distance	0·917	0·855	0·797
Longer distance	0·907	0·824	0·760

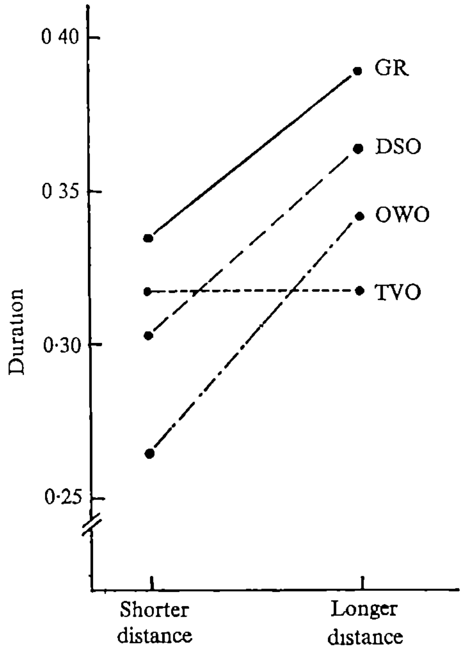


Fig. 2. Mean duration of recorded ECs (expressed as a percentage of the total time).

observers. In both conditions of distance, while the DSO's agreement was found very high, a slightly lower agreement was observed for TVOs and a still lower one for OWOs ($F = 6{\cdot}08$, d.f. $= 2, 44$, $P < 0{\cdot}005$). In fact, as shown by the use of Duncan's test, only the differences between DSOs and OWOs reached a significant level.

No significant differences were observed for the frequency of ECs recorded by the different judges. But there was a significant effect of distance on the duration of recorded EC ($F = 4{\cdot}02$, d.f. $= 1, 60$, $P < 0{\cdot}05$). With an increase of distance (Fig. 2), longer records of EC tended to appear among GRs, DSOs, and OWOs. This effect was absent for TVOs. However, the *post hoc* analysis by Duncan's test showed the difference between the shorter and longer distance to be significant only for OWOs.

DISCUSSION

The frequency of eye-directed gaze in this study probably differed from that of a natural interaction. However, as stressed by von Cranach & Ellring (1973), no independent data exist on the true distribution of sender gazes in social sit uations Another important difference between the above reported experimental situation and normal conditions of interaction was the complete absence of movements by the senders except those of the eyes. This last characteristic probably made eye-contacts harder to judge because of the lack of accompanying signals. These limitations being underlined, the data show that for two conditions of distance, a gaze-receiver, a direct-sight observer seated near the former and an observer using a TV monitoring system were statistically equivalent in their accuracy for recording the eye-gaze of a sender. This leads to two conclusions.

First, it appears that an observer who is external to the dyadic interaction, is able to detect eye-directed gazes in this interaction with a validity equivalent to the gaze's receiver himself. In this respect, the data also show a high level of agreement, i.e. reliability, between such an observer and the receiver. Moreover, an increase of the distance separating the sender from the receiver and direct-sight observer failed to result in a significant drop of their respective judgement's validity coefficients, and of their common level of agreement. These data seem to be more encouraging than those reported in the review by von Cranach & Ellring (1973) who concluded that generally the observers discriminate eye-directed gazes much worse than do the receivers. In the present study, however, an optimal position was given to the direct-observer since he was nearly on the sender–receiver axis at the same distance from the sender as the receiver, while in the studies reported by von Cranach & Ellring, the observer was placed at right angles to the centre of the sender–receiver axis. The sender's head position was mentioned by these authors as a main factor in the recognition of gaze direction. It probably explains the difference between their results and the present ones.

A second conclusion from the above data concerns the concealed observers. When the observer of the visual behaviour of an interaction must be concealed, the use of a zoom lens TV system seems a satisfying solution, the judgements performed by means of such a system being statistically equivalent to those performed by the gaze-receiver and by the direct-observer. This was not at all the case for the currently used one-way mirror system: records performed through a mirror were significantly less accurate than those performed by the other judges. Moreover, for the agreement with the gaze-receiver, the mirror-observers were significantly lower than the direct-observers. The data then support Stephenson & Rutter's (1970) opinion that 'work on eye-contact using observers behind one-way screens is of dubious value' and confirm their suggestion that 'the most accurate method at present available is to record from film taken through a zoom lens'. One might allege that the 2 ft distance which in this study separated the mirror-observer from the gaze-receiver and from the direct-observer could explain the poorer efficiency of the first. However, in the longer distance condition, the gaze-receiver and the direct-observer were seated 9 ft from the gaze-sender, while in the shorter distance, the mirror-observer was seated 8 ft from him. Nevertheless, the validity of the mirror-observer's records was significantly lower in this shorter condition than that of the gaze-receiver and direct-observer in the longer one. The only remaining explanation becomes the mirror itself. Through a one-way screen, the considerable reduction of transmitted light results in reduced visibility. Therein lies the probable cause of the mirror-observer's poorer efficiency.

Another fact in the present experiment was the tendency among gaze-receivers, direct-observers and mirror-observers to record longer EC periods in the longer distance condition than in the shorter one. While this tendency was significant beyond the 0·05 level only for the mirror-observers, the data are however in favour of the TV observational system with which no distance effect was found. Since Argyle & Dean's (1965) article on 'Eye-contact, distance and affiliation', the effect of distance on visual behaviour is well known and has stimulated a good number of studies (e.g. Goldberg *et al.* 1969; White, Hegarty & Beasley, 1970; Stephenson & Rutter, 1970; Argyle, 1970; Argyle & Ingham, 1972). Two kinds of facts are now established in that field. First, when distance increases between really interacting partners, a parallel increase in length of looks at partners' eyes is observed (Stephenson *et al.* 1972; Knight *et al.* 1973; Russo, 1975; Coutts & Schneider, 1975). Second, even when the gaze sender is acting in a standardized way throughout different conditions of distance, an overestimation of his eye-directed gazes occurs under longer distance conditions in the receiver's records as in those of the observers (von Cranach *et al.* 1968; Krüger & Hückstedt, 1969; Knight *et al.* 1973), a fact which is consistent with the results reported here. In the study by Knight *et al.* (1973) the overestimation of the sender's eye-directed gazes at longer distance was present among observers while their distance from the sender was held constant throughout the experiment, the different conditions of distance being only due to variations of the receiver's position. This leads to the hypothesis that subjects tend to be influenced in their recordings of visual behaviour by what they expect to happen in a real-life situation, or, by 'culturally defined expectations for appropriate eye-contact at different interaction distances' (Knight *et al.* 1973). The visual frame of reference seems to be of importance in the determination of such expectations effects since, in the present data, these effects were completely absent only for TV observers who, whatever the condition of distance, received through the zoom lens the same image of the sender's face alone. Future research should consider further this question of expectations about non-verbal behaviour in the interaction.

References

Anstis, S. M., Mayhew, J. E. & Morley, T. (1969). The perception of where a face or television 'portrait' is looking. *Am. J. Psychol.* **82**, 474–489.

Argyle, M. (1970). Eye-contact and distance: a reply to Stephenson and Rutter. *Br. J. Psychol.* **61**, 395–396.

Argyle, M. & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry* **28**, 289–304.

Argyle, M. & Ingham, R. (1972). Gaze, mutual-gaze, and proximity. *Semiotica* **6**, 32–49.

Cline, M. (1967). The perception of where a person is looking. *Am. J. Psychol.* **80**, 41–50.

Cohen, J. (1960). A cofficient of agreement for nominal scales. *Educ. psychol. Measur.* **20**, 37–46.

Coutts, L. M. & Schneider, F. W. (1975). Visual behavior in an unfocused interaction as a function of sex and distance. *J. exp. soc. Psychol.* **11**, 64–77.

Cranach, M. von, Hückstedt, B., Schmid, R. & Vogel, M. V. (1968). Unpublished experiments. Max-Planck-Institut für Psychiatrie, Munich (quoted in von Cranach & Ellring, 1973).

CRANACH, M. VON & ELLRING, J. H. (1973). Problems in the recognition of gaze direction. In M. von Cranach & I. Vine (eds), *Social Communication and Movement*, pp. 419–443. New York: Academic Press.

EFRAN, J. S. (1968). Looking for approval: Effects on visual behavior of approbation from persons differing in importance. *J. Person. soc. Psychol.* **10**, 21–25.

ELLRING, J. H. & CRANACH, M. VON (1972). Processes of learning in the recognition of eye-signals. *Europ. J. soc. Psychol.* **2**, 33–43.

EXLINE, R. V. (1963). Explorations in the process of person perception: visual interaction in relation to competition, sex, and the need of affiliation. *J. Personality* **31**, 1–20.

GIBSON, J. J. & PICK, A. D. (1963). Perception of another person's looking behavior. *Am. J. Psychol.* **76**, 386–394.

GOLDBERG, G. N., KIESLER, C. A. & COLLINS, B. E. (1969). Visual behavior and face-to-face distance during interaction. *Sociometry* **32**, 43–53.

KENDON, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychol.* **26**, 22–63.

KNIGHT, D. J., LANGMEYER, D. & LUNDGREN, D. C. (1973). Eye-contact, distance and affiliation: the role of observer bias. *Sociometry* **36**, 390–401.

KRÜGER, K. & HUCKSTEDT, B. (1969). Die Beurteilung von Blickrichtungen. *Z. exp. angew. Psychol.* **16**, 452–472.

LIGHT, R. J. (1971). Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol. Bull.* **76**, 365–377.

MEHRABIAN, A. & WILLIAMS, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *J. Person. soc. Psychol.* **13**, 37–58.

RUSSO, N. F. (1975). Eye-contact, interpersonal distance, and the equilibrium theory. *J. Person. soc. Psychol.* **31**, 497–502.

RUTTER, D. R. & STEPHENSON, G. M. (1972). Visual interaction in a group of schizophrenic and depressive patients. *Br. J. soc. clin. Psychol.* **11**, 57–65.

STEPHENSON, G. M. & RUTTER, D. R. (1970). Eye-contact, distance and affiliation: a reevaluation. *Br. J. Psychol.* **61**, 385–393.

STEPHENSON, G. M., RUTTER, D. R. & DORE, S. R. (1972). Visual interaction and distance. *Br. J. Psychol.* **64**, 251–257.

STRONGMAN, K. T. & CHAMPNESS, B. G. (1968). Dominance hierarchies and conflict in eye-contact. *Acta Psychol.* **28**, 376–386.

VINE, I. (1971). Judgment of direction of gaze: an interpretation of discrepant results. *Br. J. soc. clin. Psychol.* **10**, 320–331.

WHITE, J. H., HEGARTY, J. R. & BEASLEY, N. A. (1970). Eye contact and observers bias: a research note. *Br. J. Psychol.* **61**, 271–273.

(*Manuscript received* 19 *May* 1975; *revised manuscript received* 8 *September* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 515–516
Printed in Great Britain

A COMMENT ON MEDDIS' 'A SIMPLE TWO-GROUP TEST FOR MATCHED SCORES WITH UNEQUAL CELL FREQUENCIES'

By R. N. SYKES
Department of Psychology, University College, Swansea

Meddis (1975) reports what is claimed to be a new distribution-free test for a randomized blocks design frequently encountered in psychological research in which subjects are observed under two conditions and various numbers of observations are obtained in each condition for each subject. The test he suggests involves computing a z score based on the normal approximation to the distribution of the sum of the ranks given to all the observations in one of the experimental conditions, following separate ranking for each subject over both conditions.

The test is, in fact, identical to the rank-sum test with grouped data described by Wilcoxon (1947) who gives exact tables for the test for the case of $k = 2(1)7$ blocks (subjects) and $n = 2(1)7$ observations per treatment per block. Wilcoxon also describes how the normal approximation can be used, and how the test can be applied to the case with different numbers of observations in each block.

The original formulation of the test was for the case in which subjects in a two independent groups experimental design fall naturally into blocks in which the variance is less than over the observations as a whole. It is attractive to generalize this case to the case in which each block consists of a single subject with several observations from that subject in each experimental condition. However, the application of the rank-sum test to the observations from a single subject may violate the assumption of the independence of observations required by the rank-sum test, and great caution must be exercised in applying this test to the within-subjects design. In the example quoted by Meddis, if, for argument's sake, waking a subject from sleep with a sound had the effect of increasing the probability of future stimuli being incorporated in a dream and of increasing the threshold for arousal, a difference in arousal thresholds for stimuli incorporated into dream content and stimuli not so incorporated could be found which would be a function of a sequential effect only.

Bradley (1968, pp. 115–117) discusses the relative merits of adopting the procedure described by Wilcoxon (1947) compared with that involving obtaining a single observation from each subject (block), in each condition. The disadvantage involved in the use of the rank-sum test for grouped data is that ranking within blocks (subjects) results in a shortening of the ranking scale with some loss of ordinal information. Against this is the advantage which accrues when there is little overlap between the scores of different blocks (subjects).

A discussion of the efficiency of the test is given by Hodges & Lehmann (1962) and Noether (1963), and van Elteren (1960) discusses methods of optimization.

REFERENCES

BRADLEY, J. V. (1968). *Distribution-free Statistical Tests*. Englewood Cliffs, N.J.: Prentice-Hall.

VAN ELTHEREN, P. H. (1960). On the combination of independent two sample tests of Wilcoxon. *Bull. Inst. Int. Stat.* **37**, 351–361.

HODGES, J. L. & LEHMANN, E. L. (1962). Rank methods for combination of independent experiments in analysis of variance. *Ann. Math. Stat.* **33**, 482–492.

MEDDIS, R. (1975). A simple two-group test for matched scores with unequal cell frequencies. *Br. J. Psychol.* **66**, 225–227.

NOETHER, G. E. (1963). Efficiency of the Wilcoxon two-sample statistic for randomized blocks. *J. Am. Stat. Ass.* **58**, 894–898.

WILCOXON, F. (1947). Probability tables for individual comparisons by ranking methods. *Biometrics* **3**, 119–122.

(*Manuscript received* 13 *November* 1975)

Br. J. Psychol. (1976), **67**, 4, *pp.* 517–518
Printed in Great Britain

A CRITICAL NOTE ON QUINTON AND FELLOWS' OBSERVATION OF REASONING STRATEGIES

By J. ST B. T. EVANS
School of Behavioural and Social Science, Plymouth Polytechnic

Quinton & Fellows (1975), in reporting an experimental study of three-term series reasoning, make the following conclusion: '...we feel the present study has demonstrated the value of obtaining introspective information from subjects concerning their own thinking processes. Too often, in efforts to be objective, researchers have relied upon making inferences about underlying processes based on differences between subjects' performances on two tasks or two conditions. What we have shown is that subjects have a great deal to contribute to our understanding of their thinking processes if only we are prepared to ask them.'

This conclusion is in need of critical examination, not only with regard to the interpretation of their particular experiment, but also with regard to the widespread use of a naive assumption that introspective reports are a direct indication of processes underlying behaviour.

Before looking specifically at the study of Quinton & Fellows, let us examine what is generally meant by the term 'strategy'. In the context of problem solving it is always used to refer to a series of operations which culminate in the observed behavioural response. In addition it is often used to refer to a conscious thought process which implies both an *intentionality* and an *awareness* on the part of the subject. Certainly those who deduce strategies from introspective data, such as Quinton & Fellows, must be regarding both of the processes described above as being one and the same thing.

The objection to this will not be made on the hackneyed grounds that introspective reports are unreliable and inadequate, although we must not lose sight of the vast accumulation of evidence that this is true in many situations. More constructively, Wason & Evans (1975) have rejected a unitary concept of thought process and proposed a dual process model which resolves many puzzling and paradoxical results reported in the literature on problem solving and reasoning (see also Evans & Wason, 1976).

The first kind of process, which we shall call a Type 1 process, is that which underlies behaviour and can only be deduced by the use of the kind of experimental procedures referred to in somewhat disparaging terms in Quinton & Fellows' conclusion. The conscious thought process, or Type 2 process, is obtainable through introspection. Wason & Evans have shown experimentally that the two processes may operate in quite a different manner and serve quite different functions. For example, a Type 2 process may rationalize the behavioural outcome of a Type 1 process of which a subject is quite unaware. It is characteristic of Type 2 processes that the subject constructs an explanation of his behaviour which is quite consistent with what he has done. This leads to the illusion, from which experimental psychologists are not exempt, that he both intended and was aware of the process underlying his behaviour.

Let us now examine the evidence on which Quinton & Fellows base their strong conclusion. In a first stage of 'strategy discovery' subjects are given problems to solve and encouraged at intervals to verbalize their thought processes and explain how they are solving the problems. On the basis of this, Quinton & Fellows classify the strategies in two main groups, referred to as 'thinking' and 'perceptual' strategies. The last block of problems given to each subject is timed and they discover that those presumed to use 'perceptual' strategies are significantly quicker.

Thus far we have evidence only of a correlational relationship, i.e. subjects who report a certain kind of strategy also tend to solve the problems more quickly, which in no sense indicates a causal relationship between strategy and performance. A weak attempt to demonstrate causality is made by Quinton & Fellows who train two subjects with a 'perceptual' strategy both of whom produced very quick times. What should, of course, be done is to compare two groups trained by a naive experimenter to use two of the different strategies observed and predict a difference in solution time.

If the experiment recommended above were carried out with positive results, it could only indicate, at best, that in this particular situation a Type 2 process is being used to produce observed behaviour. Such a possibility is not excluded by Wason & Evans (1975). If it has happened with Quinton & Fellows' subjects this might be a result of their procedure, with the use of constant demand for verbalization of thought processes from the subject. As Neisser (1963) has pointed out, the act of verbalization may well change the nature of the thought process being studied. In any case it seems very hard to justify the dangerously naive conclusion of Quinton & Fellows which, if taken seriously, might lead us to abandon experimental psychology and regress to the methods of nineteenth century introspectionists.

REFERENCES

Evans, J. St B. T. & Wason, P. C. (1976). Rationalization in a reasoning task. *Br. J. Psychol.* (in Press.)

Neisser, U. (1963). The multiplicity of thought. *Br. J. Psychol.* **54**, 1–14.

Quinton, G. & Fellows, B. J. (1975). 'Perceptual' strategies in the solving of three-term series problems. *Br. J. Psychol.* **66**, 69–78.

Wason, P. C. & Evans, J. St B. T. (1975). Dual processes in reasoning? *Cognition* **3**, 141–154.

(*Manuscript received* 1 *September* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 519–520
Printed in Great Britain

THE ROLE OF INTROSPECTION IN PROBLEM-SOLVING RESEARCH: A REPLY TO EVANS

BY BRIAN J. FELLOWS
Department of Social Studies, Portsmouth Polytechnic

The basic question at issue in this short note is whether by looking inward or reflecting upon our own mental processes while we are solving problems we can obtain any valid information about the processes underlying our behaviour. As long as man has been reflecting upon his mental activity, and this extends at least as far back as written records are preserved, he has assumed that by doing so he can discover some of the determinants of his own behaviour. It is this assumption that Evans (1976) is questioning in his critical note on the Quinton & Fellows (1975) study. The aim of this present paper is to suggest that in some circumstances this assumption is quite reasonable and to question it is self-contradictory.

What Quinton & Fellows did was ask their subjects to solve a large number of three-term series problems and to encourage them to talk about the way they were tackling them. Their subjects reported using two main types of strategy. In the 'thinking' type the subject said that he tried to represent the terms in some sort of meaningful manner, often in a spatial image. In the 'perceptual' type the subject reported using certain invariant display characteristics of the problem to arrive at a short-cut solution. Quinton & Fellows found that subjects tended to move from 'thinking' to 'perceptual' strategies as more problems were solved and that solution times became shorter as that happened. They concluded that this demonstrated the value of asking subjects about their own thought processes rather than relying on making inferences about these processes from the way subjects perform under different experimental conditions.

What Evans does is to argue that the processes underlying problem-solving behaviour can only be discovered by the experimental–inferential procedure (his Type 1 processes) and dismisses subject reports as attempts by the subject to rationalize his own behaviour. Thus, the picture we have of a subject is of a passive and relatively confused observer of his own physical behaviour which is determined by factors outside of his control and awareness. When asked by the experimenter to report on his thought processes the hapless subject constructs an explanation which he thinks is consistent with what he has done and which he hopes will impress the experimenter.

To deal first with the point about self-contradiction. If Evans dismisses the subject's reports as rationalizations then logically he must also dismiss his own explanations in the same way. Are the experimenter's explanations any more than constructions which he thinks are consistent with the subjects' behaviour and which he hopes will impress other experimenters? But the view I wish to support in this paper is to break down the artificial distinction between 'the experimenter' who is all wise and 'the subject' who is a fool. Rather they are in the quest for knowledge together and any experimenter who fails to acknowledge the contribution a subject can make, particularly in studies of problem solving though not only there, is liable to get a warped though consistent view of what is going on. Let me try to illustrate this by tracing the early development of the Quinton & Fellows study.

All the studies we could locate on the three-term problem were of the experimental–inferential kind. The resulting theories all emphasized the importance of setting up some sort of internal representation of the premises, often of a visual kind (see Wason & Johnson-Laird, 1972, chapter 9 for a summary). These theories seemed very reasonable until we began asking subjects to tell us what they were doing when solving this sort of problem. Some of their reports fitted the theories, but others definitely did not. A number of subjects reported strategies which we did not expect and had little reason to expect from studying the experimental literature. It was not until later that we came across the work of Wood (1969), Wood & Shotter (1973) and Wood, Shotter & Godden (1974) who had also found non-representational strategies reported by subjects in this type of problem. Wood went on to affirm the existence of these strategies in an 'objective' manner by asking unexpected supplementary questions which could only be answered by subjects who had formed a representation of the premises. We could have done something similar by introducing the unexpected indeterminate or negative problem. But these tests do little more than affirm what we already know from subjects' introspections. If we had done these tests before asking, as Evans would recommend, then we might finish up with a more complex and consistent representational theory but one which did not match up to what the subject said he was doing. In this case whose judgement should we trust? Can we say that the experimenter is right irrespective of what the subject says?

Evans concludes by arguing that, if taken seriously, the Quinton & Fellows conclusion might lead to the abandonment of experimental psychology. If by 'experimental psychology' he means the treatment of human subjects as if they were nonverbal animals, which is not an uncommon interpretation of the term, then by all means this is the implication. But this is not the only meaning of the term. To experiment is also to do something in order to find out something which is not known. And if we have to seek the advice of our subjects in order to do that then quaint ideological notions should not stand in our way. As Wason & Johnson-Laird (1972, p. 5) have commented introspection '...as a method for revealing the processes of reasoning is a very useful auxiliary technique'.

REFERENCES

EVANS, J. ST B. T. (1976). A critical note on Quinton and Fellows' observation of reasoning strategies. *Br. J. Psychol.* 67, 517–518.

QUINTON, G. & FELLOWS, B. J. (1975). 'Perceptual' strategies in the solving of three-term series problems. *Br. J. Psychol.* 66, 69–78.

WASON, P. C. & JOHNSON-LAIRD, P. N. (1972). *Psychology of Reasoning*. London: Batsford.

WOOD, D. J. (1969). Approach to the study of human reasoning. *Nature, Lond.* 223, 101–102.

WOOD, D. J. & SHOTTER, J. D. (1973). A preliminary study of distinctive features in problem solving. *Q. Jl exp. Psychol.* 25, 504–510.

WOOD, D. J., SHOTTER, J. D. & GODDEN, D. (1974). An investigation of the relationships between problem solving strategies, representation and memory. *Q. Jl exp. Psychol.* 26, 252–257.

(*Manuscript received* 21 *November* 1975)

Br. J. Psychol. (1976), 67, 4, *pp*. 521–528
Printed in Great Britain

REHEARSAL SPAN VERSUS MEMORY SPAN: WHAT DOES THIS COMPARISON TELL US ABOUT VERBAL MEMORY?

BY LAWRENCE PORTER
University of Oxford, Department of Experimental Psychology

In the rehearsal span (RS) task the subject is presented with a series of items such that each new item is presented only after he has repeated all the preceding items in the correct order. Contrary to the results reported by Dalrymple-Alford (1967), it was found (*a*) that the maximum number of digits, on average, a subject could repeat back under these conditions (the rehearsal span) was significantly greater than the conventional memory span (MS); and (*b*) that recall of early items in the series benefited from the cumulative rehearsal in the RS task. In addition, ordered free recall of RS series was superior to that of MS series. Presentation rate had no significant effect on the magnitude of any of the above differences between performance in the RS task and that in the MS task. These results do not support previous conclusions, based on the failure to observe any difference between MS and RS, about the role of rehearsal in short-term verbal memory.

Theories of verbal memory fall into two categories: those which allow that rote rehearsal, *per se*, can strengthen memory traces and those which do not. An example of the former type is the 'dual store' theory of Atkinson & Shiffrin (1968) who suggest that there are two functionally separate memory stores, the short-term store (STS) and the long-term store (LTS), and that 'rehearsal serves the purpose of increasing the strength built up in a long-term store both by increasing the length of stay in STS (during which time a trace is built up in LTS) and by giving coding and other strategies time to operate' (p. 111). A recent alternative to the dual store model, one which falls into the second category, is the 'levels of analysis' theory advanced by Craik & Lockhart (1972) and Craik (1973). They propose that memory is a by-product of the activity of a perceptual system which can encode stimuli at various levels of analysis, ranging from crude physical attributes to sophisticated semantic attributes. They maintain that all the phenomena which have previously been held to support the dual store notion have derived from the fact that information encoded at different levels shows different forgetting characteristics. A corollary to this theory is that rehearsal, as such, does not strengthen the memory trace but simply maintains it at a given level of analysis. Any such strengthening arises only through deeper analysis of the material to be remembered.

One way of furnishing evidence for the levels-of-analysis view of rehearsal would be to show that, in conditions where the subject is unlikely to have opportunity for 'deeper' encoding, the effect of rote rehearsal on recall of a verbal item is no greater than that of a single presentation occurring at the time of the last rehearsal of the item. An experiment which purported to have demonstrated this phenomenon was carried out by Dalrymple-Alford (1967). He compared the cumulative rehearsal span (which will subsequently be referred to simply as the rehearsal span or RS) with the memory span (MS). In the RS task the subject is presented with a series of digits such that each new digit is presented only after he has repeated all the preceding digits in the correct order. The maximum number of digits, on average, a subject can repeat back without error is his rehearsal span. Dalrymple-Alford (1967) found (*a*) that no significant difference existed between RS and MS (with rate of presentation and rate of recall the same for both conditions); and (*b*) that the subject's first error in an RS series was equally likely to occur at any serial position. He argued that, having failed to find any effect of the substantial number of rehearsals accorded to early list items in the RS condition, 'repetition restores but does not further strengthen memory traces' (p. 63). Tulving & Madigan (1970) submit that Dalrymple-Alford's results show rehearsal to be, under some conditions, 'totally ineffective' in producing learning greater than that produced by a single presentation (p. 448). The aims of the present study are (1) to show that Dalrymple-Alford's inferences from his data were unwarranted; and (2) to investigate more fully the differences, if any, between MS and RS.

With respect to the former aim, a statistically non-significant difference between MS and RS provides little foundation for conclusions about the effectiveness of rote rehearsal. In fact, three of Dalrymple-Alford's ten subjects showed RSs significantly greater than their MSs. (Admittedly he used an overpowerful statistical test, justifiable when negative results are claimed.) He suggests that, for these particular subjects, the presentation rate may not have been sufficiently brisk to prevent recoding in the RS condition. The hypothesis implied here, that increasing presentation rate reduces the advantages, if any, of RS over MS, is tested in Expt. II below.

Dalrymple-Alford (1967) rightly places more significance on the evidence provided by serial position effects in RS. However, it is incorrect to suggest that equiprobability of breakdown on all serial positions of an RS list implies that all the items are of equal strength in memory. Suppose the probability of making an error on any given item is a constant, P, and is independent of the probability of making an error on any other item. Then the probability of making the *first* error on the ith item is P times the probability that no error has already occurred on a previous item in that series, i.e. $P(1-P)^{i-1}$. For series of a given length, k, on those trials on which an error does occur, the probability of that first error occurring on the ith item is $P(1-P)^{i-1}/(1-P^k)$. Clearly, if all items are equally and independently recallable, the proportion of first errors should be a decreasing function of serial position. Conversely, by this simple model, observation of equiprobable breakdown at all serial positions indicates that early items have been given a higher probability of recall than later items. It is not at all clear how any 'equal trace strength' model, compatible with an observed rectangular distribution of the serial position of the first error, could be devised. Further, it is trivial to compare the proportions of RS errors in various serial positions with any theoretical function in the absence of similar data from a control MS condition. Large serial position effects are commonly found in short-term memory experiments. It is therefore impossible to assess the effects of rehearsal (with respect to serial position) in the RS task without an MS baseline. In Dalrymple-Alford's (1967) study there were insufficient MS trials to provide reliable data on serial position of breakdown in the MS condition.

Two experiments are reported here. The first is a replication of Dalrymple-Alford's (1967) experiment. In the second, RS is compared with MS at three presentation rates, one equal to, one faster than and one slower than that used by Dalrymple-Alford. In addition to recording span length and serial position of first error, an attempt is made in Expt. II to achieve more sensitive measurement by observing, on each trial, the subject's ordered free recall of the series of items on which he makes his first error.

EXPERIMENT I

In Expt. I the procedure followed that of Dalrymple-Alford (1967) as closely as was practicable using a computer rather than a memory drum to present the stimuli.

Method

Subjects

Ten young adults, graduate or undergraduate members of the University of Oxford, served as paid volunteers. Each subject was tested individually in one session.

Materials

For the RS condition, 17 12-digit lists were assembled, the individual items being selected (with replacement) from the digits 1–9 inclusive. Selection was random within the constraints that there were no ascending or descending runs of three or more consecutive digits (e.g. 543) and no digit was repeated without there being at least two other intervening digits. For the MS condition, six lists of 4, 5, 6, 7, 8, 9, 10, 11 and 12 digits were assembled in the manner just described.

Procedure

A Digital Equipment Corporation 'Linc Eight' computer was used to control the presentation of stimuli on an oscilloscope screen. Before each trial in both the RS and MS conditions the display 'READY' was presented for two seconds and the first digit of that trial followed one second after the termination of the 'READY' display.

In the RS condition the first digit in the series was followed by a cue (a display of three asterisks), the second digit by two consecutive cues, the third by three cues, and so on. The subject read each digit aloud as it was presented and then attempted to repeat all the digits he had read so far, in the correct order and in pace with occurrence of the cue. All the stimuli were 0·5 sec in duration and the interstimulus interval (ISI) was 0·5 sec, giving a presentation rate of one item per second. A trial was terminated on the subject's first error. The experimenter noted the length of the RS (the maximum number of items the subject had been able to 'carry' without error) and the serial position at which the first error had occurred. In the RS condition there was one practice trial followed by 16 experimental trials.

During the course of each session the subject's MS was estimated by his mean score on five trials, two of these (plus one practice trial) occurring before and three after the RS trials. With the computer program used, this ordering of the two conditions was more convenient than that used by Dalrymple-Alford (1967), who presented four MS trials interspersed with three blocks of five RS trials (i.e. a total of 15 RS and 4 MS trials). In the MS condition four digits were presented consecutively. The subject read these aloud and then attempted to repeat them in the correct order and in pace with four consecutive cues (the cue being again a display of three asterisks). There followed immediately a series of five new digits followed by five cues. The subject continued to read out and repeat back successive series of digits, each series being one item longer than its predecessor. As in the RS condition, all stimuli were 0·5 sec in duration and the ISI was 0·5 sec. A trial was terminated on the subject's first error. The experimenter noted the length of the MS for that trial, i.e. the length of the longest string of digits the subject had managed to repeat correctly. There were insufficient MS trials in this replication of Dalrymple-Alford's (1967) experiment to render the data on the serial position of breakdown of any value.

Results

Each subject's mean RS and MS was computed. A t test for related samples carried out on these data revealed the overall mean rehearsal span (7·06 items) to be significantly greater than the overall memory span (6·62 items): ($t = 2{\cdot}61$, d.f. $= 9$, $P < 0{\cdot}05$, two-tailed). Since the MS trials were not evenly distributed throughout

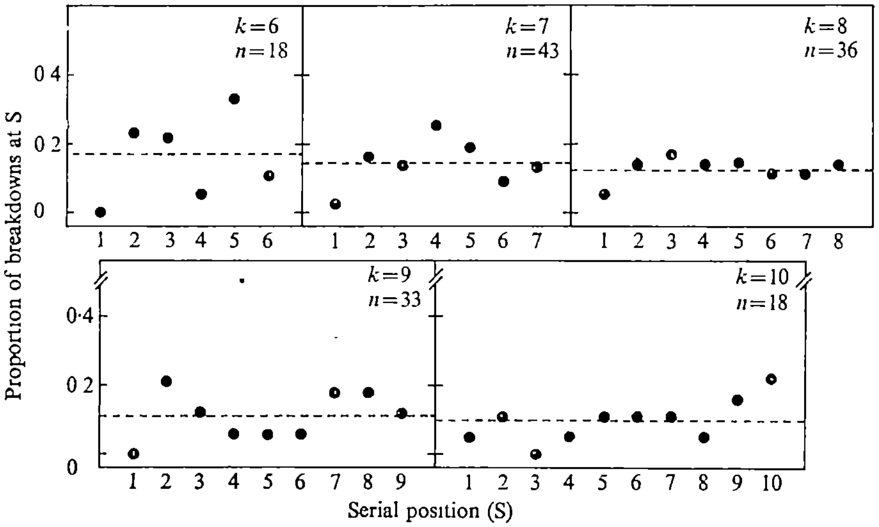


Fig. 1. Proportion of breakdowns on the RS task as a function of serial position at which breakdown occurred; k = number of digits to be repeated; n = number of observations on which proportions are based.

the experimental session, this difference may be due to fatigue effects on the final MS trials. However, a t test comparing MS on trials at the beginning of the session with that on trials at the end revealed no significant effect of fatigue (average MS at the beginning = 6·45 items and average MS at the end = 6·73 items; t = 0·99, d.f. = 9, $P > 0·3$, two-tailed).

The distribution of serial positions at which recall broke down in the RS condition is shown in Fig. 1 for lists of 6, 7, 8, 9 and 10 items. Data from trials on which the subjects made an error on a series of fewer than 5 or more than 10 items have been omitted since these trials occurred infrequently (8 per cent of the total). For direct comparison with the data reported by Dalrymple-Alford (1967), the present results were analysed following the procedure used by him. For each of the five distributions shown in Fig. 1, the Kolmogorov–Smirnov one-sample test failed to reveal departure from equiprobability of breakdown at all serial positions. (The theoretical functions representing equiprobable breakdown at all serial positions are the dashed lines in Fig. 1.) This result is in agreement with Dalrymple-Alford's (1967), but as has been noted above, cannot provide unambiguous evidence about the effects of cumulative rehearsal in RS.

Discussion

That RS significantly exceeds MS is contrary to the findings of Dalrymple-Alford (1967). It is possible, however, that presentation rate was slow enough for the present subjects to engage in recoding strategies, such as grouping, in the RS condition. It is also possible that RS performance was superior to MS performance as a result of the greater amount of practice subjects received in the former condition. In Expt. II, in addition to variable presentation rates, there were equal numbers of RS and MS trials. Equating the numbers of RS and MS trials also allows the isolation of those serial position effects in recall which are due solely to the cumulative rehearsal present in the former, but absent in the latter task.

EXPERIMENT II

Method

Subjects

These were eight young adults, graduate or undergraduate members of the Universities of Oxford or Durham, who served as paid volunteers. Each subject was tested individually in four sessions.

Materials

In the RS condition the digit lists were similar to those in Expt. I except that they were extended to 16 items in length. For the MS condition, the digit lists were assembled in identical fashion to those in Expt. I.

Procedure

The stimuli were presented using the same apparatus as in Expt. I. Every trial was preceded by a 'READY' display, of two seconds duration, which terminated one second before the appearance of the first digit in the series.

The structure of both the RS and the MS trials was identical to that in Expt. I. In the present experiment, however, there was a total of 30 RS and 30 MS trials. In one-third of these (i.e. 10 RS and 10 MS trials) the presentation rate was two items per second (stimulus duration = 0·25 sec, ISI = 0·25 sec). This was labelled the 'fast' condition. In another third of the trials, the 'medium' condition, presentation rate was one item per second (stimulus duration = 0·5 sec, ISI = 0·5 sec). In the remaining third of the trials, the 'slow' condition, presentation rate was 0·5 items per second (stimulus duration = 1 sec, ISI = 1 sec).

Each subject completed four sessions separated by intervals of approximately 24 hours; one session in the fast condition, one in the medium condition and two sessions (each comprised of 5 MS plus 5 RS trials) in the slow condition. Four subjects, chosen at random, always completed all the MS trials before completing all the RS trials in each session, while four underwent these conditions in the reverse order. The order in which each subject underwent the fast, medium and slow conditions was chosen at random, the two slow sessions always occurring consecutively. Within a session each condition (MS or RS) was immediately preceded by one appropriate practice trial.

A trial was terminated on the subject's first error. The experimenter noted the serial position of the first error and recorded the length of the immediately preceding (error free) series. On hearing the signal terminating the trial, the subject wrote down as many of the digits as he could remember, attempting to preserve their correct order but writing them down in any order he chose (ordered free recall). Since a response was to be scored as correct only if it was the correct item in the correct serial position, it was necessary for the subject to make exactly the same number of responses as there were items in the series he was attempting to recall. Thus the signal to terminate the trial was in the form of the spoken word 'Stop' followed by a spoken number, e.g. 'Stop: seven', where the number referred to the length of the series the subject was to attempt to recall on that trial. He was instructed to make exactly that number of responses, with omissions indicated by dashes.

Results

The effects of two independent variables, span type (MS or RS) and presentation rate (fast, medium or slow), on three dependent variables, span length, serial position of breakdown and ordered free recall, were analysed.

Span length

The span length for each trial was taken as the maximum number of items the subject managed to repeat back, in the correct order, without error. A three-way

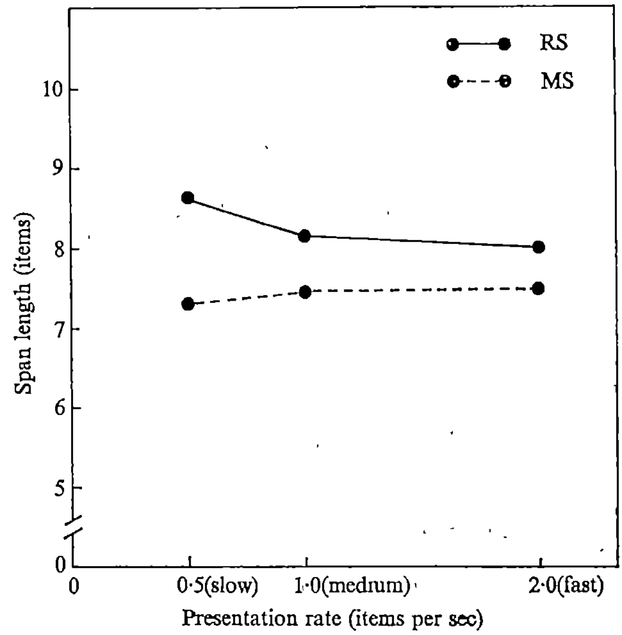


Fig. 2. Span length as a function of presentation rate; ●---●, MS; ●—●, RS.

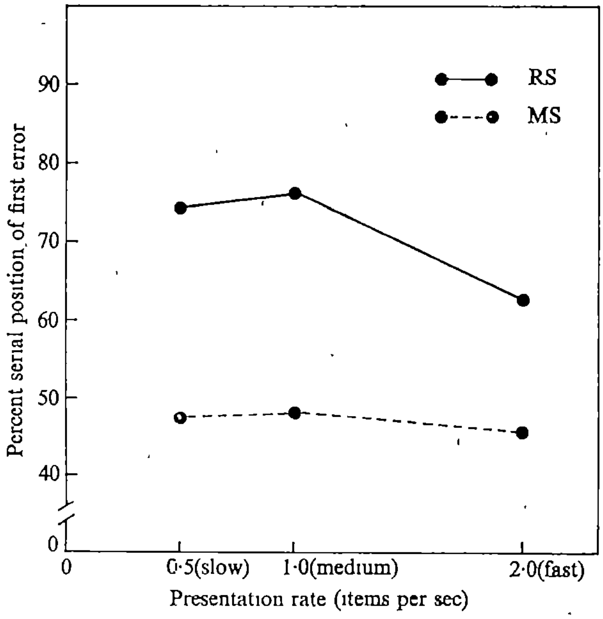


Fig. 3. Serial position of first error (expressed as a percentage of critical list length) as a function of presentation rate; ●---●, MS; ●—●, RS.

factorial analysis of variance was carried out, the factors being span-type (2 levels), presentation rate (3 levels) and subjects (8 levels, random), there being 10 replications in each cell. The RS (8·25 items on average) was found to be significantly greater ($F = 39{\cdot}46$, d.f. = 1, 7, $P < 0{\cdot}001$) than the MS (7·40 items on average). No other main effect or interaction reached significance. Preplanned t tests showed that RS was significantly greater than MS at each of the presentation rates ($P < 0{\cdot}025$, $P < 0{\cdot}05$ and $P < 0{\cdot}01$ for fast, medium and slow rates respectively). The data on span length are shown in Fig. 2.

Serial position of breakdown

For each trial, the serial position of breakdown, expressed as a proportion of critical list length for that trial, was computed. An analysis of variance similar to that for span length was performed on the resulting data. The main effect of span-type was significant ($F = 125{\cdot}07$, d.f. = 1, 7, $P < 0{\cdot}001$). In the MS condition the first error occurred, on average, 46·9 per cent of the way through the list while, in the RS condition, it occurred 71·0 per cent of the way through the list. The effect of presentation rate was significant ($F = 6{\cdot}19$, d.f. = 2, 14, $P < 0{\cdot}025$), breakdown occurring earlier in the fast condition than in the slow or medium conditions (see Fig. 3). The interaction between span-type and presentation rate was not significant. Preplanned t tests showed that the first error occurred at a significantly earlier serial position under MS than under RS at each of the presentation rates ($P < 0{\cdot}01$, $P < 0{\cdot}001$ and $P < 0{\cdot}001$ for fast, medium and slow rates respectively).

Ordered free recall

In the subject's written output a response was scored correct only if it was the correct item in the correct serial position of occurrence in the critical list. An analysis of variance similar to that for span length was performed on these data. The main effect of span-type was significant (average RS per cent recall = 71; average MS per cent recall = 55; $F = 16{\cdot}22$, d.f. = 1, 7, $P < 0{\cdot}001$). The effect of presentation rate was significant (average per cent recall for fast, medium and slow rates of presentation = 57, 61 and 70 respectively; $F = 5{\cdot}63$, d.f. = 2, 14, $P < 0{\cdot}01$). The interaction between span-type and presentation rate was not significant.

Discussion

Significant differences between RS and MS were found on all the observed measures. Thus the comparison of rehearsal span with memory span has provided no evidence for the abandonment of the hypothesis that rote rehearsal, *per se*, consolidates or strengthens information in memory. The results of Dalrymple-Alford (1967) supported the second category of theory mentioned above, that exemplified by the 'levels of analysis' theory of Craik & Lockhart (1972). The contrary results reported here, while they remove that support, can be predicted by either type of theory. Any increase in recall with cumulative rehearsal could result from increased opportunity for recoding rather than from the rehearsal itself.

The failure to find any interaction between presentation rate and span-type weakens the 'recoding' interpretation of the present results, since, at fast rates, there should be less time for subjects to engage in recoding activity. However, the absence of these interactions must be interpreted with caution. Inspection of Figs 2 and 3 reveals a tendency for the predicted interactions to occur. This tendency is borne out by the magnitude of the F values concerned ($F = 3{\cdot}00$ and $2{\cdot}41$ respectively, d.f. = 2, 14). Also, there may be a limit to the extent to which recoding is reduced by increasing presentation rate, particularly in tasks, such as those reported here, where the subject is allowed to vocalize the items to be remembered, thus producing patterns of intonation. Grouping by intonation pattern is an important means of recoding series of digits presented at fast rates (Murray, 1967). This kind of recoding might be reduced by having the subject make an irrelevant verbal response between successive presentations or rehearsals of items.

I am grateful to the Northern Ireland Department of Education for financial support, to Dr N. C. Waugh and Dr G. A. Gelade for comments on drafts of the manuscript and to S. Fearnley, who wrote the computer program.

References

Atkinson, R. C. & Shiffrin, R. M. (1968). Human Memory: a proposed system and its control processes. In K. W. Spence & J. T. Spence (eds), *The Psychology of Learning and Motivation*, vol. 2. New York: Academic Press.

Dalrymple-Alford, E. C. (1967). Repetition and immediate memory. *Br. J. Psychol.* **58**, 63–67.

Craik, F. I. M. (1973). A 'levels of analysis' view of memory. In P. Pliner, L. Krames & T. M. Alloway (eds), *Communication and Affect, Language and Thought*. New York: Academic Press.

Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. verb. Learn. verb. Behav.* **11**, 671–684.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Can. J. Psychol.* **21**, 263–276.

Tulving, E. & Madigan, S. A. (1970). Memory and verbal learning. *Ann. Rev. Psychol.* **21**, 437–484.

(*Manuscript received* 4 *July* 1975; *revised manuscript received* 11 *November* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 529–535.
Printed in Great Britain

THE EFFECT OF EMBEDDED WORDS IN A BRIEF VISUAL DISPLAY

BY DAVID E. SOMEKH
*Formerly: Department of Psychology, Bedford College**

In a replication and extension of an experiment by Eagle, Wolitzky & Klein (1966), subjects wrote brief stories describing an Object Relations Test card following exposure to a 7 × 7 letter matrix in which were embedded either neutral words or emotive words. Of six groups of subjects who participated, two groups were exposed to the matrices for 10 sec ('supraliminal control condition'), and the remaining four groups for 1 sec ('subliminal condition').

Independent judges were able to distinguish, to a significant degree, between the stories of subjects who were exposed to emotive words and those exposed to neutral words *under the 'subliminal condition'*; they could not distinguish, however, between the stories of subjects exposed for the longer duration ('supraliminal condition') to emotive and neutral words respectively.

A crude content analysis yielded significantly more unpleasant words in stories from the subgroups exposed to emotive rather than neutral words, but again only under the very brief exposure condition, the difference being absent in supraliminal control groups. The relevance of these findings to previous experiments is discussed.

There is evidence to show that stimuli which may be unperceived at the time of their occurrence can be retrieved under conditions which favour the operation of primary thought processes (Fisher, 1960; Silverman & Silverman, 1964; Dixon, 1971).

Eagle *et al.* (1966) looked specifically at the effects of a part of the stimulus display of which the observer was not aware. They used an embedded figure (a duck in a tree) which produced significantly more 'duck related' items in subsequent imagining of a nature scene than did a control figure, even though subjects were unaware of the source of their imagery. They concluded '... connotative aspects of the ground, although apparently unperceived, may become manifest in non-perceptual response modes such as imagery' (p. 839). Of importance to selective attention models is the level of analysis which such an effect implies.

It was decided to use a similar technique to Eagle *et al.* but to take their study a stage further in two ways. First, a within awareness control group was used to attempt to show a qualitative difference between the two situations and second, embedded emotive *words* were used, with words of very similar structure but different meaning as controls.

METHOD

Subjects were tested in groups, subjects within each group being arbitrarily divided into two equal subgroups labelled 'A' and 'B' respectively. Group A (experimental group) were briefly exposed to a matrix of letters with two emotive words, FUCK and PAIN, embedded in it. Group B (control group) were shown a very similar matrix except that the embedded words were DUCK and RAIN. Both groups then wrote down as much of the matrices as they could remember and then the first ten words that came into their heads. The whole group was then shown an Object Relations Test card and asked to write a brief story about it. Stories were collected, typed and randomized and blind judges with knowledge of the procedure were asked to assign

---

* Now at The Maudsley Hospital, Denmark Hill, London, S.E.5.

R	X	H	G	B	R	D
T	E*	Y	J	Q	L	R
Z	S	V	P	A	I	N
W	D	B	M	T	Q	Z
J	H	R	Y	G	V	S
Q	T	X	F	U	C	K
L	M	Z	D	J	B	M

P	X	H	G	B	P	F
Y	O*	T	J	Q	L	H
Z	S	V	R	A	I	N
W	F	B	M	T	Q	Z
J	H	P	Y	G	V	S
Q	T	X	D	U	C	K
L	M	Z	F	J	B	M

Fig. 1. Stimuli used (white upper case letters on black background except for two red letters marked thus *).

the stories, on the basis of their content, to group A or group B. It was argued that the emotive words would affect the content of the stories given by group A sufficiently for the judges to be able to distinguish between stories from the two subgroups to a significant extent.

In a pilot experiment a 6 × 6 letter matrix, with a red letter in one of the four quadrants and a three-letter neutral word in differing relation to it, had been shown to 40 subjects tachistoscopically. Subjects pressed a thumb-switch which illuminated a card bearing the matrix and started a stop-clock. The subject then reported the initial target (red letter) verbally and pressed the thumb-switch when he or she found the secondary target (a neutral word – reporting it verbally as a check), thus stopping the clock to register total search time. The fastest time from 240 trials (12 × 20 subjects), when the word and letter were not actually adjacent, was 1·35 sec (mean time 2·91 sec). It was therefore argued that with a 7 × 7 letter matrix and the subject not aware that words were present in the matrix, a 1 sec exposure, even if repeated, would make it virtually impossible for subjects to report the embedded words. On the other hand in view of Yarbus's finding (1967) that during a 5 sec period the observer of a complex picture changed points of fixation 18 times, it is likely under these conditions that subjects would have performed several saccades during the 1 sec period and therefore an image of the area of the matrix containing the embedded words would have fallen on the retina.

Materials

The stimuli are shown in Fig. 1. These and Object Relations Test cards B2 and B3 were mounted as 2 × 2 in. slides and projected on to a screen at an average distance of 12 ft 6 in. from the projector. A 300 watt Leisegang magazine-loading projector was used with a Prontor shutter connected via a voltage stabilizer to a Devices digitimer. Subjects were given protocol sheets which merely had a 7 × 7 open matrix in the top left corner and below, the numbers 1–10, for the association exercise.

Subjects

One hundred and twenty-four subjects (42 male, 82 female), mainly undergraduates (some of them in psychology) and sixth-formers.

Procedure

Subjects were tested in groups of between 16 and 24. Alternate subjects were handed protocol sheets and told 'you belong to group A (or B), please write your letter in the right-hand corner of the sheet'. Having divided subjects into equal groups and checked that no subject was red colour-blind, the experimenter said, 'I am now going to show you a matrix of letters. Somewhere in the matrix will be a red letter which is your primary target. As soon as you have found the red letter, try and see as much of the rest of the matrix as you can. I'll show you the matrix very briefly twice, separated by a few seconds. Immediately afterwards I want you to write down as much of the matrix as you can remember in the space provided. Write as many letters as you can, even if you're not quite sure they're right, beginning with the red letter. Now, here's the difficult bit, for which I need your cooperation. I'm going to show the matrix to group A first and while I'm doing this I want group B to look down at the desk or anyway, promise not to look. As soon as group A are trying to write down what they saw, I shall show group B a similar matrix. So the procedure will be: group A look at the board, group B look down. Then, group A try and write down what they saw while group B look at the board. Any questions?'

At the command, 'group A look at the board, group B look down', group A were shown one of the two matrices twice for 1·0 sec, separated by a gap of 3 sec. Approximately 15 sec after group A were first shown their matrix, group B were shown the other of the two matrices. After approximately 2 min subjects were told, 'If you've written down as much as you can remember, now write down the first ten words which come into your head, in the space provided. When you have finished, turn your sheet over.'

When all subjects had finished, they were told, 'I am now going to show you a picture. After you have seen it I want you to write a short story, about 6–8 lines, about what you think is happening – that is, I'd like you to write about two lines on how you think the situation came about, four to six lines on what is happening now and about two lines on what you think will happen next'. Subjects were shown one of the two ORT pictures for 10 sec and then wrote their stories. Protocols were collected and subjects were asked what they thought the experiment was about and its true nature was explained.

Four groups of subjects, totalling 80 subjects, 40 experimental and 40 control, were tested in this manner. Order of presentation of neutral and emotive matrices and use of ORT card B2 or B3 were randomized across groups. On checking the protocols, without exception all subjects had written down the correct red letter (either an E or an O) corresponding to their status as group A or B. No subject had written down either of the embedded words. On questioning, no subject was aware that embedded words had been present.

In the 'supraliminal' session, two groups of subjects, numbering 20 and 24, were tested. The procedure was exactly the same except that each exposure of the matrices was for 10 sec instead of 1 sec. By questioning and from scrutiny of the protocols it was established that a total of ten subjects of the 44 had seen and reported one or both of the stimulus words. PAIN was reported five times, FUCK four times, RAIN three times and DUCK not at all.

RESULTS

The stories for each group were randomized and typed and were sent to five judges with details of procedure. The judges' task was to suggest, by reference to story content, which came from the experimental subgroup and which from the control subgroup. The five judges for each group of stories were two naive judges (NJ), usually psychology postgraduates with no detailed knowledge of the Object Relations Test technique; two expert judges (EJ), psychoanalysts with knowledge of the ORT, and one semi-experienced judge (SJ), usually a psychology lecturer with some knowledge of the ORT. A total of 14 judges were used for the 30 judgements (five judges for each of the six experimental groups).

Table 1 shows the number of correct judgements by each of the five judges for each group. To test for significance, the results of each group of stories were organized into a contingency table according to the number of judges making a correct judgement on a particular story from group A or group B. Each table was then analysed for significance, using Kendall's tau (with correction for ties) and the result expressed as a z score. A cumulative z score over successive groups was then obtained.

Over a total of 80 subjects in the 'subliminal condition', judges were able to distinguish between the two groups to a significant extent, $z_c = 2{\cdot}027$ (one-tailed, $P < 0{\cdot}025$). When matrices were shown to subjects for a longer period, judges were unable to distinguish between stories coming from the 44 subjects, $z_c = 0{\cdot}071$ (n.s.).

These findings lend tentative support to the suggestion that stories given after brief exposure to hidden emotive words are different from those following exposure to hidden neutral words and that longer exposure to the words reduces the effect. The exact nature of the difference is difficult to determine as each judge used his own criterion for assessment. There was, nevertheless, general agreement that stories

Table 1. *Summary of story judgements*

(Number correctly assigned to group A or B)

	Number of subjects in A+B	Judges					Kendall's S	z	Cumulative z_0
		NJ_1	NJ_2	EJ_1	EJ_2	SJ			
'Subliminal' sessions									
Group 1	16	8	10	10	10	8	19	1·003	1·003
Group 2	24	12	16	16	16	12	47	1·370	1·680
Group 3	16	8	10	10	10	10	23	1·217	2·073
Group 4	24	10	14	12	16	12	16	0·464	2·027 (one-tailed, $P < 0{\cdot}025$)
'Supraliminal' sessions									
Group 5	20	10	10	8	8	14	−2	−0·076	−0·076
Group 6	24	14	12	14	12	10	6	0·175	0·071 (n.s.)

following brief exposure to the words FUCK and PAIN showed more disturbance, whether violent themes, uncertainty, conflict or loss, than stories following exposure to the words DUCK and RAIN. In an attempt to provide a more objective demonstration of differences between the two groups of stories, a crude content analysis of the 124 stories was performed.

Content analysis

From 124 stories, average length 80 words, approximately 1500 different words were extracted. These were randomized and presented to seven independent judges who were asked to rate them as pleasant, neutral or unpleasant. It was decided to take as a criterion a minimum of three out of seven judgements of U or P, with U and P judgements cancelling each other out if they co-occurred. In fact, of the 349 words qualifying as U or P words, only 17 (< 5 per cent) were not uniformly judged by a minimum of three judges (the others judging the word as neutral).

The stories were then scored for content of unpleasant or pleasant words. Overall, story lengths and proportion of non-neutral words were similar across groups. However, in the 'subliminal' sessions, stories from the experimental group A tended to have more unpleasant than pleasant words, the reverse being true for stories from control group B. The only exception was in group 2 where stories from group A also had more pleasant than unpleasant words. Overall, in experimental condition stories 45 per cent (112) of non-neutral words were pleasant, as compared to 60 per cent (153) from control condition stories. In the 'supraliminal' sessions, the proportion of pleasant to unpleasant words did not differ for stories from group A and B, the proportion of pleasant words being 56 per cent (58) and 52 per cent (71) respectively.

In order to assess the significance of these findings, following a suggestion from Mr J. D. Valentine of Bedford College, each story was assigned a 'difference score' which represented the proportion of pleasant to unpleasant words, as follows:

$$\text{index} = \frac{P-U}{P+U},$$

and could therefore vary between −1 and +1.

Table 2. *Content analysis: mean pleasant/unpleasant index scores*

(S.D.S in parentheses)

	Group A (experimental)	Group B (control)	(n)	U	P
'Subliminal' sessions					
Group 1	−0·20 (0·77)	+0·31 (0·81)	8 × 2	20	0·12
Group 2	+0·16 (0·58)	+0·12 (0·56)	12 × 2	70	0·55
Group 3	−0·55 (0·57)	+0·13 (0·72)	8 × 2	11	0·01
Group 4	−0·18 (0·71)	+0·30 (0·75)	12 × 2	40	0·04
'Supraliminal' sessions					
Group 5	+0·06 (0·39)	−0·04 (0·53)	10 × 2	55	0·65
Group 6	+0·06 (0·68)	+0·03 (0·80)	12 × 2	55	0·16

Table 3. *Frequency of U and P words expressed as a percentage of the total number of words in the stories*

	Experimental condition		Control condition	
	% U	% P	% U	% P
'Subliminal' sessions	4·24	3·41	3·30	5·00
'Supraliminal' sessions	2·66	3·35	3·50	3·82

Within each of the sessions indices from experimental and control stories were compared, using the Mann–Whitney U test. Table 2 shows the mean index values for each group, the U values comparing indices within each session and associated one-tailed probabilities. Probabilities were summed across sessions, using the formula

$$X^2_{2k} = -2\,\Sigma_1^k \log_e P_i,$$

which yielded, for the 'subliminal' sessions,

$$X^2 \text{ (d.f. 8)} = 20{\cdot}48 \ (P < 0{\cdot}01)$$

and for the 'supraliminal' sessions

$$X^2 \text{ (d.f. 4)} = 4{\cdot}48 \ (P > 0{\cdot}3)$$

Table 3 shows subliminal and supraliminal session totals expressed as a percentage of the total number of words in the stories.

It appears from the table that there is a tendency under the subliminal condition for emotive stimuli such as FUCK and PAIN to produce stories with more unpleasant words, while neutral stimuli such as DUCK and RAIN produce stories with more pleasant words. When the words approach awareness, i.e. some subjects are able to report them, the percentage of unpleasant words in the experimental condition and of pleasant words in the control condition both decreased; the percentage of pleasant words in the experimental condition and of unpleasant words in the control condition remained unchanged, however. The larger fall in the percentage of unpleasant words in experimental (group A) supraliminal stories might tentatively be ascribed to a defensive response, suggesting a quite different handling of the emotive stimuli when they are on the borders of awareness than when subjects are totally unaware of them.

DISCUSSION

The findings appear to confirm those of many previous workers (see reviews by Fisher, 1960 and Dixon, 1971) that marginal or incidental stimuli, like subliminal stimuli, may be registered without awareness and can be recovered in a variety of modes of experience and behaviour – in this case, symbolically transferred as imagery in a story. The use of suitable controls makes explanations of the phenomenon in terms of partial cues of 'fleeting' perceptions improbable. As Eagle *et al.* comment, perceptions too fleeting to be reported would seem to be operationally indistinguishable from stimulus registrations without awareness.

It can be argued that some of the subjects in the 'supraliminal' sessions should have been responding to the stimuli in the same way as subjects in the 'subliminal' sessions, as they were unable to report the stimuli subsequently. The z scores for judgements of groups 5 and 6 are, however, both lower than that for any of the 'subliminal' groups, and it would appear that the incidental stimuli had a much weaker effect when presented near the awareness threshold. This is consistent with the findings of Paul & Fisher (1959) and others, that the subliminal effect increases as stimuli are presented further below threshold.

In the present experiment a considerable number of factors contribute to error variance. To give two examples, the pleasant–unpleasant measure is very rough and ready; and secondly, the word RAIN may have a strong effect with some subjects when followed by ORT card B2 which shows shadowy figures apparently sheltering under a tree outside a large house (Phillipson, 1973). Rather than detract from the findings, this observation suggests that the possibility for a more sophisticated methodology which may contribute better to an understanding of the mechanism underlying the effect and facilitate a considerable reduction in error variance. For example, if the individual subjects were tested with the ORT before the experiment, to obtain baseline profiles, the specific effect of the incidental stimuli could then be assessed and this procedure would enable use of objective scoring methods, which are available (Phillipson, 1973).

The implications for theories of selective attention are twofold. First, consistent with the results of previous studies the findings tend to support the view that incidental stimuli may be recovered or have an effect after a period exceeding that normally associated with short-term memory. Second, the difference between experimental (group A) and control (group B) conditions appears to depend on subjects discriminating words of very similar structure but different meaning. This supports findings from previous experiments (e.g. Somekh & Wilding, 1973), which suggest that incidental or 'irrelevant' stimuli are processed to a high level of analysis.

This experiment was supported by a grant from the Social Science Research Council as part of the work for a doctorial thesis (London University). The author would like to thank Dr J. Sandler for helpful suggestions and Dr J. M. Wilding for encouragement and advice. The author would also like to thank the postgraduates who acted as judges and (in alphabetical order) Mrs P. Antonis, Mr V. B. Kantor, Dr M. Lawlor, Mr H. Phillipson and Mrs A. M. Sandler. Special thanks to Mrs S. Fierz.

REFERENCES

DIXON, N. F. (1971). *Subliminal Perception*. London: McGraw-Hill.

EAGLE, M., WOLITZKY, D. I. & KLEIN, G. S. (1966). Imagery: effect of a concealed figure in a stimulus. *Science, N.Y.* **151**, 837–839.

FISHER, C. (1960). Preconscious stimulation in dreams, associations and images. *Monogr. 7: Psychol. Issues* **2**, 1–40.

PAUL, I. H. & FISHER, C. (1959). Subliminal visual stimulation: a study of its influence on subsequent images and dreams. *J. nerv. ment. Dis.* **129**, 315–340.

PHILLIPSON, H. (1973). Personal communication.

SILVERMAN, L. H. & SILVERMAN, D. K. (1964). A clinical–experimental approach to the study of subliminal stimulation. *J. abnorm. soc. Psychol.* **69**, 158–172.

SOMEKH, D. E. & WILDING, J. M. (1973). Perception without awareness in a dichoptic viewing situation. *Br. J. Psychol.* **64**, 339–349.

YARBUS, A. L. (1967). *Eye-movements and Vision*. New York: Plenum Press.

(*Manuscript received* 29 *July* 1974; *revised manuscript received* 16 *September* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 537–552
Printed in Great Britain

THE COMPONENTS OF THE POGGENDORFF ILLUSION

BY R. H. DAY AND R. G. DICKINSON
Department of Psychology, Monash University

An explanation of the Poggendorff misalignment illusion in terms of three basic components, the horizontal–vertical, longitudinal–transverse and obtuse angle effects, is proposed. It is argued that these effects may either singly or jointly give rise to an apparent elongation–contraction of the space between the aligned elements with consequent change in apparent oblique direction. Experimental evidence for the involvement of the three components in various combinations and for modification of perceived direction between the transversals is presented. It is shown that the illusion in various forms of the Poggendorff figure can be accounted for by the involvement of one or more of the three components, each of which can be demonstrated independently of the illusion. It is suggested that other classical illusions, including the Müller-Lyer, might also be analysed in terms of basic components.

In its conventional form the Poggendorff illusion is an apparent misalignment of two collinear, oblique transversals separated by a pair of parallels (Fig. 1*c*). However, the effect occurs in a wide range of figures including those in which the transversals are upright and the parallels oblique (Tolanski, 1964; Day & Dickinson, 1975), one transversal is replaced by an aligned dot (Weintraub & Krantz, 1971; Krantz & Weintraub, 1973), one or both parallels are removed (Weintraub & Krantz, 1971; Goldstein & Weintraub, 1972; Day, 1973*a*; Curthoys, Wenderoth & Harris, 1975) and in which the obliques form only acute or obtuse angles with the parallels (Day, 1973*b*; Imai, 1973). Selective reviews of early and recent investigations can be found in Gillam (1971), Hotopf & Ollerearnshaw (1972*a*, *b*), Day (1973*a*, *b*, 1974), and a more extensive review and discussion in Robinson (1972).

Hering (1861), Helmholtz (1886) and Wundt (1893) attributed the misalignment illusion to apparent angular displacement of the transversals towards a right angle with the parallels, a perceptual effect which Hotopf & Ollerearnshaw (1972*a*, *b*) called the regression to right angles tendency. This effect has been recently confirmed and measured by Blakemore, Carpenter & Georgeson (1970) and Hotopf & Ollerearnshaw (1972*b*), the latter finding, however, that it accounts for only about half the misalignment effect. Later Hotopf, Ollerearnshaw & Brown (1974) reasoned that three angular effects are involved in the apparent displacement of the transversals and reported experimental data in support of this view. They concluded that the whole of the misalignment illusion can be accounted for by these effects.

The validity of an explanation based on angular displacement of the transversals is seriously questioned by the results of a recent experiment (Day & Dickinson, 1975). A circular field of parallel lines or a single median line, both pointing straight ahead, either did not change in apparent orientation or changed only very slightly *away* from a right angle when an oblique, solid rectangle which interrupted the lines was introduced at the centre of the field. Nevertheless, marked apparent misalignments of the collinear interrupted lines were reported. These and recent confirmatory data* have led to the conclusion that the Poggendorff illusion occurs strongly without any apparent regression of the transversals to right angles.

---

* The experiment can be criticized on the ground that the long sagittal transversals which extended fully across the visual field were too perceptually stable to exhibit apparent regression to right angles with the oblique rectangle. However, essentially the same results were obtained in a more recent but unreported experiment with short (20 mm) transversals.

The purpose of this paper is to propose an alternative explanation of the Poggendorff illusion in terms of differences in apparent extent in the space *between* the aligned elements, with consequent change in apparent oblique direction. Apparent extent can be shown to be relatively elongated in one direction and relatively contracted in that at right angles to it. Three independent components, the horizontal–vertical (HV), longitudinal–transverse (LT) and the obtuse angle (OA) effects contribute either singly or jointly to this perceptual distortion. In consequence of this perceptual elongation apparent oblique direction across the open space is modified and the aligned elements apparently misaligned. In short, a linear distortion of perceptual space results in a change in apparent oblique direction. Thus the misalignment effect is secondary and has its origin in one or more of three basic illusions, the HV, LT and OA effects.

The paper is divided into five sections as follows: description of and evidence for the involvement of the HV, LT and OA effects; evidence for change in apparent oblique direction in the space between the aligned elements, and absence of change in the direction of the transversals; discussion of the process involved in apparent misalignment; consideration of the extent to which the occurrence of the illusion in various figures can be accounted for by one or more of the components; and consideration of the broader implications of the explanation.

Numerous hitherto unreported experiments in support of the explanation are described briefly. Since the apparatus, psychophysical procedure, experimental design and subject population have been fully described in four earlier reports (Day, 1973*a*, *b*, 1974; Day & Dickinson, 1975) further detailed description is unnecessary here. In all the experiments described the subject was required to adjust a moveable feature, usually a line or a dot, four times per condition, twice in each direction, with the order of directions balanced. Each subject participated in each condition with the order of the conditions either balanced or randomized across the group. The score for each subject in each condition was the mean of the four adjustments. Any departures from this standard procedure are described. Details of stimulus figures, number of subjects, results, and the outcomes of statistical tests are given for each experiment.

INVOLVEMENT OF THE HORIZONTAL–VERTICAL, LONGITUDINAL–TRANSVERSE AND OBTUSE ANGLE EFFECTS IN THE POGGENDORFF ILLUSION

The HV, LT and OA effects have in common the relative elongation of visual extent in one direction and its apparent contraction in that at right angles to it. As will be shown, the magnitude of this perceptual distortion varies between components which occur singly or in various combinations in particular figures.

The horizontal–vertical (HV) component

A visual extent which corresponds with the vertical meridian of the eye appears about 2–6 per cent greater than the same extent falling along the horizontal meridian (Finger & Spelt, 1947; Künnapas, 1955; Begelman & Steinfield, 1967; Avery & Day, 1969). While this effect is usually demonstrated with an L-figure, it also occurs when

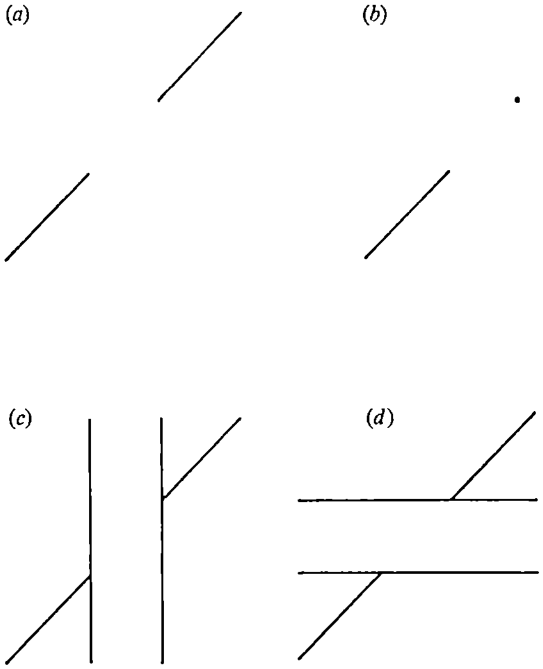


Fig. 1. Demonstrations of the contribution of the horizontal–vertical component to apparent misalignment. In (*a*) and (*b*) this component alone is involved. The component contributes positively in (*c*) and negatively in (*d*).

a single line is presented alone and compared with an implicit or 'imagined' standard (Avery, 1969). That the effect is a function of the visual meridian stimulated rather than of sheer verticality is shown by the reversal of the illusion when either the subject or the figure is turned through 90° (Avery & Day, 1969). The 'pure' HV effect is distinguishable from the bisection effect but the two occur together in a T-figure (Finger & Spelt, 1947; Künnapas, 1955).

Since vertical extent is apparently greater than objectively equal horizontal extent two collinear oblique lines would be expected to appear slightly misaligned. Maximum misalignment would be expected when the vertical and horizontal extents between the parallels are equal, i.e. when the lines are at 45°, and zero effect when both are either upright or horizontal. A small but significant apparent misalignment occurs in a figure consisting simply of two collinear oblique lines arranged as in Fig. 1(*a*) (Wundt, 1893; Sanford, 1898; Goldstein & Weintraub, 1972; Day, 1973*a*). The effect reaches a maximum at 45° when perceptual misalignment is about one-third that for the conventional figure shown in Fig. 1(*c*) (Day, 1973*a*) but is insignificant when the lines are vertical or horizontal.

Direct support for the involvement of the HV effect in the Poggendorff illusion is provided by the usually small but nevertheless consistent reduction in the Poggendorff illusion when the normally upright figure (Fig. 1*c*) is turned through 90° so that the parallels are horizontal as in Fig. 1(*d*) (Leibowitz & Toffey, 1966; Weintraub & Krantz, 1971). In the vertical position the HV effect would be expected to combine with other components and increase the illusion and in the horizontal position to be

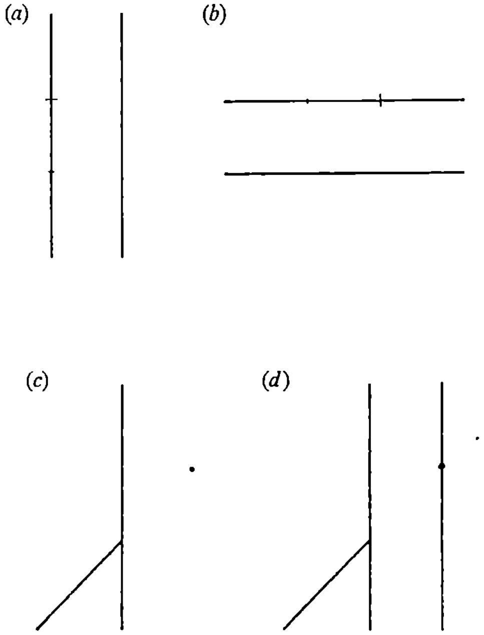


Fig. 2. Method used to measure the longitudinal–transverse effect. In (*a*) and (*b*) the distance between the marks on the parallels was adjusted to match that between the parallels. The component is present in (*d*) with two parallels upright but not in (*c*) with one upright.

opposed to them and thus decrease the illusion. Therefore, half the difference between the mean misalignment in the upright and horizontal orientations would serve as an indication of the HV contribution.

In an experiment using the figures shown in Fig. 1 (*c*) and 1 (*d*) the parallels were 100 mm long and 30 mm apart, standard dimensions for the parallel line figure throughout, and the 45° obliques 50 mm long and 42 mm apart. A group of 14 subjects adjusted the right (Fig. 1*c*) and upper (Fig. 1*d*) obliques to apparent alignment using the procedures described above. Mean misalignment for the upright condition was 6·85 mm (S.D. = 2·57) and for the horizontal 5·27 (2·51), a difference of 1·58 mm which was significant ($t = 3{\cdot}08$, d.f. = 13, $P < 0{\cdot}01$).

The longitudinal–transverse (LT) component

Piaget & Denis-Prinzhorn (1953) found that the long side of a rectangular figure is overestimated and the short side underestimated relative to a comparison stimulus, an effect they called the illusion of rectangles (see also Piaget, 1969). Similar observations were made earlier by Wundt (1893) and Sanford (1898) in both rectangular and parallel line figures. Two experiments have confirmed that in a parallel line figure apparent longitudinal extent is greater than an objectively equal transverse extent, an effect which will be called here the longitudinal–transverse effect.

A group of 12 subjects moved a 2 mm mark along a parallel so that the distance between it and a 5 mm mark, also on the parallel, appeared the same as that between the parallels. The figure in which the parallels were 100 mm long and 30 mm apart with the 2 mm and 5 mm marks symmetrically arranged on one of them is shown in Fig. 2(*a*). Mean underestimation of transverse extent by longitudinal extent was 4·23 mm (S.D. = 1·41), indicating that a longitudinal extent of 25·77 mm was apparently equal to a transverse extent of 30 mm. Of course, with the figure upright, the HV component would have contributed to this difference.

In order to separate the HV and LT components 12 new subjects carried out the same task with the figure upright and horizontal as in Fig. 2(*a*) and 2(*b*). As in the alignment experiment described above (Fig. 1*c* and 1*d*) it was expected that the HV component would add to underestimation of transverse by longitudinal extent in the upright position and reduce it in the horizontal. Therefore, half the difference between the mean underestimations in the two orientations served as an indication of the HV component. Mean underestimation for Fig. 2(*a*) was 3·50 mm (S.D. = 2·57) and for Fig. 2(*b*) −0·56 mm (2·75), i.e. a slight overestimation. Half the difference, 2·03 mm, can be taken as the HV contribution, notably greater than in the alignment experiment using two orientations.

The LT component occurs in a figure in which transverse extent is defined by the parallels. It would be expected not to contribute, or to contribute only slightly, to apparent misalignment in a figure with a single upright. To test this prediction 12 subjects adjusted a dot to apparent alignment with a transversal in the figure shown in Fig. 2(*c*) and 2(*d*), the dimensions of which were the same as in Fig. 1(*c*). It was expected that in Fig. 2(*c*) the HV and OA components would determine the illusion and in Fig. 2(*d*) the HV, OA and LT components. Therefore, the difference between mean misalignment in the two figures would indicate the size of the LT contribution. Mean misalignment for Fig. 2(*c*) was 5·3 mm (S.D. = 2·33) and for Fig. 2(*d*) 7·4 mm (2·88). The difference of 2·1 mm was significant (t = 3·806, d.f. = 11, P < 0·01).

The obtuse angle (OA) component

The apparent length of the arm of an obtuse angle is greater than that of an acute angle of equal physical extent. This difference was first reported by Müller-Lyer (1889) who regarded it as the basis of the illusion in his more complex 'arrow' figures. The difference was described also by Laska (1890), Jastrow (1891) and Judd (1899). Day (1974) confirmed that apparent extent was greater along the boundary of the obtuse angle than for an equal standard extent and less along that of an acute angle. It can be reasonably assumed, therefore, that visual extent along that part of a parallel forming an obtuse angle in a typical Poggendorff figure (Fig. 1*c*) is perceptually elongated.

Both Krantz & Weintraub (1973) and Tong & Weintraub (1974) have reported results which clearly demonstrate the contribution of obtuse angles to the misalignment illusion. The illusions were consistently larger in those figures in which one or more obtuse angles occurred and markedly greater when the two lines actually met obtusely at an intersection.

The contribution of the OA component was established in an experiment described in the following section, which was designed also to assess the relative contributions of the three components.

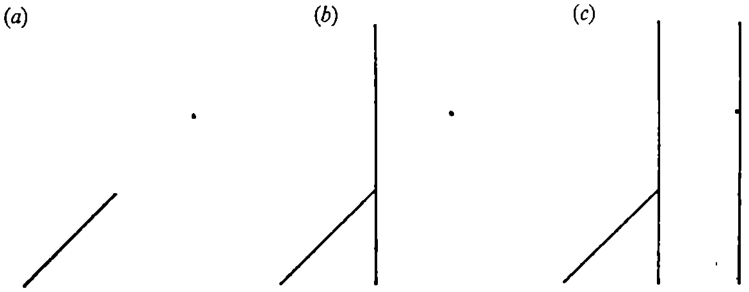


Fig. 3. Figures used to investigate the relative contributions of the horizontal–vertical (HV), longitudinal–transverse (LT) and obtuse angle (OA) components to apparent misalignment. In (a) the HV effect alone contributes, in (b) the HV and OA, and in (c) the HV, OA and LT.

Relative contributions of the three components

In the experiments described so far the size of the HV and LT components have varied. The magnitudes of the HV component were 0·79 and 1·78 mm and those of the LT component 1·72 and 2·10 mm. These differences are probably due to different tasks (alignment and matching distances) and different samples. The following experiment was designed to establish the size of the OA component and to assess the relative contributions of the three components using the same task – aligning a dot with a line – and subjects.

Twelve subjects adjusted a 1 mm dot to apparent alignment with a 50 mm, 45° oblique line in three figures. The figures which were the same as those in Figs 1 (*b*), 2 (*c*) and 2 (*d*) are shown together in Fig. 3 for convenience. It was assumed from earlier findings that in Fig. 1 (*a*) the HV component alone would contribute to apparent misalignment, in Fig. 1 (*b*) the HV and OA components, and in Fig. 1 (*c*) the HV, OA and LT components. Thus the contribution of the HV component was represented by the mean misalignment in Fig. 1 (*a*), the OA by the difference between means in Fig. 1 (*a*) and 1 (*b*), and the LT by the difference between Fig. 1 (*b*) and 1 (*c*).

The mean misalignments for Fig. 1 (*a*), 1 (*b*) and 1 (*c*) were respectively 2·50 (S.D. = 1·31), 5·43 (1·83) and 7·21 mm (1·41). An analysis of variance showed that the overall difference between means was significant ($F = 42{\cdot}79$, d.f. = 2, 22, $P < 0{\cdot}005$). In addition, two planned contrasts showed that the difference in mean misalignments between Fig. 1 (*a*) and 1 (*b*) and between Fig. 1 (*b*) and 1 (*c*) were significant ($F_{AB} = 32{\cdot}48$, d.f. = 1, 22, $P < 0{\cdot}005$; $F_{BC} = 11{\cdot}99$, d.f. = 1, 22, $P < 0{\cdot}005$). Thus the relative contributions of the HV, LT and OA components to the total illusion of 7·21 mm were 2·50, 1·78 and 2·93 mm representing 33·6, 26·0 and 40·6 per cent respectively. As implied in earlier work (Tong & Weintraub, 1974) the OA effect is the major contributor to the perceptual distortion underlying the misalignment effect, followed in order by the HV and LT effects.

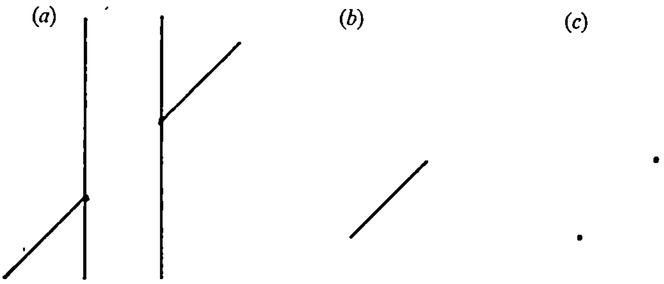


Fig. 4. Arrangement for investigation of change in apparent direction of transversal (*a* and *b*) and of oblique direction between parallels (*a* and *c*). The change was measured by comparing matches to direction in (*a*) and (*b*) and in (*a* and (*c*).

APPARENT OBLIQUE DIRECTION IN THE POGGENDORF ILLUSION

If in the figures shown in Fig. 1 visually perceived extent in one direction is relatively elongated it would be expected that apparent oblique direction would also undergo change. That is, if in the parallel line figure (Fig. 1*c*) the apparent extent of a given distance along the figure is greater than the same extent across it, then a particular oblique direction would be apparently steeper. The greater the apparent elongation of a figure the greater the change in apparent direction obliquely across it. Thus, if the HV, LT and OA effects all combined to produce marked apparent elongation, apparent oblique direction would accordingly be steeper than if only one or two of these components contributed.

Two experiments showed that apparent oblique direction in the space between the parallels is markedly steeper than it is without the parallels. In addition, the first experiment confirmed that apparent oblique direction (orientation) of a transversal is the same as it is without the figure.

Oblique direction of the transversals and the space between

A group of 24 subjects adjusted the angle of a 42 mm line at the centre of a circular horizontal field so that its orientation matched that of each of four standards. The standards were each at the centre of another field which was frontal, and were of the same size (20 cm) and at the same distance (42 cm). Thus the subject looked down to adjust the orientation of the variable and forward to view the standard. Adjustments were made by rotating the whole horizontal circular field. Subjects made four adjustments for each standard, 16 in all for each subject. Directions of rotation and starting direction were balanced across subjects.

The four 45° standards are shown in Fig. 4. The elements to be matched in Fig. 4(*a*) were the upper or lower transversal, half the subjects matching the orientation of one and half that of the other, and the oblique extent between the two intersections marked by 1 mm dots. The dots were placed at the intersections to indicate this extent clearly. Fig. 4(*b*) was a control for the transversals and consisted simply of a 45° oblique line by itself. Fig. 4(*c*) was a control for the oblique orientation between the parallels and consisted simply of two 1 mm dots arranged at the same angle and the same distance (42 mm) apart.

It was expected that if the transversals underwent apparent angular change there would be a difference between the matches to the oblique lines with and without parallels. Likewise, it was expected that if a change in orientation, i.e. oblique direction, across the parallels is the basis of the misalignment illusion, there would be a difference between the matches with and without the parallels. The only statistical tests planned, therefore, were between the two pairs of mean orientations, orientation of transversals and that between the parallels. For a change in the orientation of the transversal to account for apparent misalignment it would need to be in the direction of a greater angle than 45°, i.e. a regression to a right angle. However, for a change in orientation in Fig. 4(*c*) to account for the illusion it would need to be in the direction of a 'steeper' angle, i.e. towards a *smaller* angle with the parallels. All measurements and computation of means were made with the upright treated as 0°.

The matched orientations of the transversal and its control (Fig. 4*a* and 4*b*) were respectively 46·42° (S.D. = 2·57) and 46·59° (S.D. = 1·67), a difference of only 0·17° which was not significant ($t = 0{\cdot}41$, d.f. = 23, $P > 0{\cdot}05$). The matched orientations of the dots (Fig. 4*a* and 4*c*) were 46·16° and 42·39° respectively. The difference of 3·77° proved significant ($t = 4{\cdot}44$, d.f. = 23, $P < 0{\cdot}01$).

These results based on 24 subjects indicate that it is not a change in the apparent orientation of a transversal which is the basis of the Poggendorff illusion, but in that of the oblique angle across the figure marked by the points of intersection. In general the results are in accord with those reported earlier, which showed that sagittal transversals do not change in direction although they are noticeably misaligned when interrupted by an oblique rectangle (Day & Dickinson, 1975). Furthermore, the results are not in accord with those of Hotopf & Ollerearnshaw (1972*b*) who reported an apparent regression of the transversal angle towards a right angle. It can be noted, however, that these investigators used a different angle (30°) and a different stimulus display in which the variable was in the same field as the standard.

A methodological point

While the results of the last experiment show that apparent orientation (oblique direction) between parallels is more acute than for the same direction without parallels, the magnitude of the difference is not enough to account for apparent misalignment in the same figure. A difference of 3·77° would result in an alignment error of about 4·2 mm whereas that found in numerous experiments with the same figure (Day, 1973*a*, *b*, 1974) has consistently been of the order of 7 mm. However, it is possible that oblique direction in the control itself (Fig. 4*c*) was also affected, since the HV component would presumably have affected apparent vertical and horizontal distances between the dots. To establish more accurately the true magnitude of the change in oblique direction between the parallels an alternative technique was used in another experiment. Twenty-four subjects merely adjusted one dot of a pair, each on parallels as in Fig. 4(*a*), so that they were at an apparent angle of 45°. This angle was first shown to subjects using a 45° set square. Four adjustments were made by each subject. The mean setting was such as to indicate that an angle of 37·4° is perceptually equivalent to 45°, a difference of 7·6°. This difference between the actual and apparent angle, which is free of the error in method using the matching technique, would account for an apparent misalignment of 7·4 mm.

THE BASIS OF THE POGGENDORFF ILLUSION

Together the results of the two experiments described in the last section and those reported earlier (Day & Dickinson, 1976) show that the transversals in the Poggendorff figure do not change in apparent orientation but the space between them does. Since a transversal and its collinear extension in the free space between the aligned elements have different apparent orientations the elements must appear misaligned. It is reasonable to suppose that the point of inflexion between the oblique line and its extension occurs at the end of the line. A consequence of this change in perceived orientation is that the greater the distance between the aligned elements (or the greater the distance between the parallels in the conventional figure) the greater the magnitude of the Poggendorff illusion. This prediction was tested in three separate experiments which although conducted at different times involved identical procedures. They can conveniently be described together.

Effect of distance between aligned elements

The purpose of the three experiments was to test the prediction that apparent misalignment in the Poggendorff illusion increases as a function of separation between the aligned elements in three figures (Fig. 5*a*, 5*b* and 5*c*). Furthermore, by calculating a line of best fit for the mean misalignments it was possible to establish the point of inflexion in orientation.

Three separate groups of 12 subjects adjusted a 1 mm dot four times to apparent alignment with an oblique line at each of the three distances. The three figures used in each of the three experiments are shown in Fig. 5. The 45° oblique was 50 mm long and in Fig. 5(*b*) and 5(*c*) the upright lines were 100 mm. The three separations between line and dot in each figure were 42, 67 and 92 mm, the first being the same as that in a parallel line figure with 30 mm between the parallels and the second and third representing 25 mm steps.

Mean misalignments for the three distances in each figure together with the least squares lines of best fit are shown in Fig. 5(*a*), 5(*b*) and 5(*c*). The means for 42, 67 and 92 mm in Fig. 5(*a*) were 1·27 (S.D. = 1·58), 2·63 (3·21) and 3·24 (3·94) mm, in Fig. 5(*b*) 4·63 (1·78), 8·27 (2·91) and 11·63 (3·63) mm and in Fig. 5(*c*) 7·36 (2·88), 13·21 (4·63) and 15·88 (7·28) mm. It can be noted that the lines of best fit intersect the oblique transversals near to the ends of the former, suggesting that change in apparent orientation does in fact occur at that point. It is clear from Fig. 5 that for all figures the size of the misalignment effect increases with separation between the aligned elements. It can be noted also that, as would be expected from the involvement of the HV, LT and OA effects, the magnitude of the illusion is greatest in Fig. 5(*c*), less in Fig. 5(*b*) and least in Fig. 5(*a*). In this regard the data correspond with those described in connexion with Fig. 3.

Asymmetrical processing in the Poggendorff illusion

Krantz & Weintraub (1973) have adduced evidence in favour of asymmetry in the perceptual processing of the Poggendorff figure. They argue that the subject 'misprojects' one transversal across the space to the opposite parallel. In this sense the point of intersection of the opposite transversal and parallel would merely serve as a

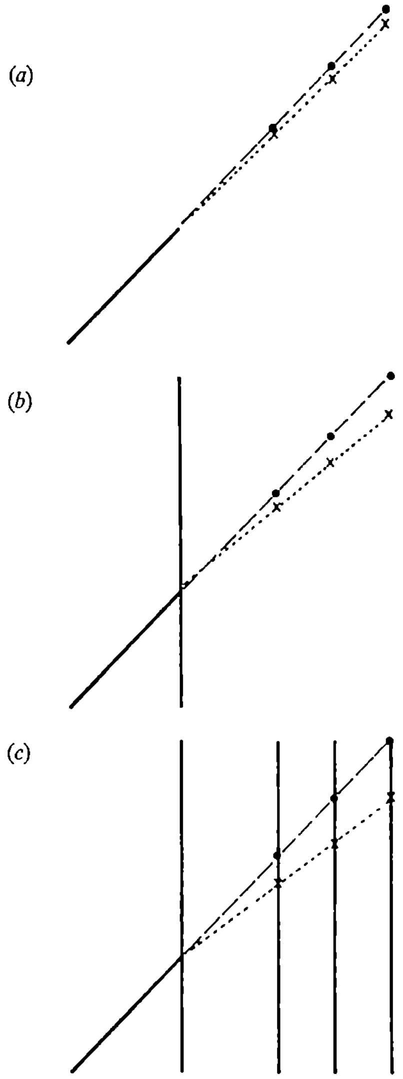


Fig. 5. Actual and apparent direction between oblique line and aligned dot. The dot was placed at three distances from the end of the oblique and adjusted to apparent alignment.

point to be aligned with the projected transversal. The experiments described in the last section support this notion. The results plotted in Fig. 5 indicate that the observer mis-projects the direction of the transversal in the space between the aligned elements. It can be presumed that the distortion of space between the parallels results in this mis-direction relative to the direction of the transversal. Asymmetry of processing in the form of mis-projections of direction across the space between the aligned elements is also supported by the observation that the illusion is constant when the aligned point of the opposite parallel is an intersection with a collinear transversal, an orthogonal transversal or simply a dot (Day, 1974). In brief, the Poggendorff task appears to involve perceptual projection of a transversal obliquely across the open space between the aligned elements; since the space is perceptually distorted mis-projection occurs.

POGGENDORFF FIGURES AND DETERMINANTS

The immediate cause of apparent misalignment is a difference between true and apparent direction between aligned elements in various figures. This modification of visual direction is a consequence of relative elongation of space due to the HV, LT and OA effects acting singly or jointly. It follows that the magnitude of the Poggendorff illusion will vary with the number of these components which occur in a particular figure and with variation in the magnitude of each component. For example, the illusion will be greater if the HV and OA components both contribute to apparent elongation than if only one of them is involved. Furthermore, if the obtuse angle is large the illusion will be greater since apparent extent is greater for large angles compared with small.

While this review cannot be exhaustive, an attempt will be made to reconcile variation in illusion magnitude with the presence or absence of the three components and with variation in the magnitude of components. While some of the interpretations of illusion magnitude and direction which follow are somewhat speculative they seem justified on the basis of data reviewed above. In any case, the interpretations are easily tested.

Poggendorff illusion figures

As pointed out above, the Poggendorff illusion occurs in a wide range of figures, a number of which are reproduced by Robinson (1972). Variation in the size of the illusion in different figures is attributable to the number of components involved, their positive or negative contributions, and the magnitude of each. For example, in the conventional figure in the upright position (Fig. 1*c*) the HV, LT and OA components jointly contribute to the effect while in the horizontal position the HV component detracts from the joint effect of the other two, as shown earlier. Furthermore, since the OA component increases with the size of the obtuse angle the illusion is greater for large obtuse angles or, as this is usually described, with smaller acute angles (Burmester, 1896; Velinsky, 1925; Weintraub & Krantz, 1971). It can be noted also that the LT effect varies directly with the length–width ratio of a figure (Piaget & Denis-Prinzhorn, 1953). So far there has been no investigation of the effect of the length–width ratio of a figure on the size of apparent misalignment. However, such an inquiry would need to control width and vary length since variation of the latter affects the illusion independently, as shown above.

Negative Poggendorff illusions

Krantz & Weintraub (1973) have reported instances of negative misalignment illusions, the largest of which occurs in a figure consisting of a line and two dots as shown in Fig. 6(*a*). A negative illusion implies a reversal of apparent elongation and contraction of space between the aligned elements. It is likely that such occurs in Fig. 6(*a*). The implied transversal between the lower left dot and the vertical line forms an acute angle, the boundary of which would be shorter than the same extent by itself. Given that the LT component is absent from this figure and assuming that decrease of apparent extent due to the implied acute angle exceeds elongation from

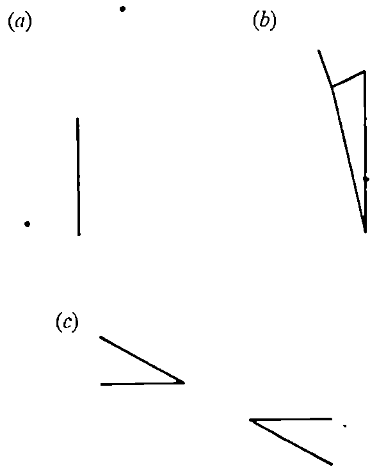


Fig. 6. Examples of negative Poggendorff illusions probably resulting from presence of implicit (*a*) or actual (*b*) acute angle boundaries. (*a*) from Krantz & Weintraub (1973); (*b*) from Pressey & Sweeney (1972).

the HV component, a negative illusion would be expected. The negative illusions reported by Krantz & Weintraub (1973) occur in figures in which shortening of apparent extent due to implied acute angle between dot and line in the absence of an LT component would be expected to occur.

A somewhat different state of affairs obtains in the figure shown in Fig. 6(*b*) in which Pressey & Dewar (1970) and Pressey & Sweeney (1972) report a negative illusion, although shortening of visual extent due to an acute angle is probably involved. The dot in Fig. 6(*b*) is apparently too high for alignment with the oblique line, not too low as would be expected from apparent elongation of the figure. The dot lies along the boundary of an acute angle which would be apparently shorter and the transversal along that of a large obtuse angle. Both elongation of the obtuse angle and shortening of the acute angle would be expected to combine to produce a distortion in the triangular figure.

Reports of negative Poggendorff illusions in acute angle figures similar to that in Fig. 6(*c*) (Hotopf & Ollerearnshaw, 1972*b*) have not been consistent (see Day, 1973*b*). Restle (1969) reported a negative illusion when the parallels were upright and the obliques formed angles of 30°. Day (1973*b*) found a negligible effect with the same figure but a small positive illusion when the angles were 45°. Possible sources of inconsistency in terms of the present explanation of the Poggendorff illusion are variation in the size of the HV component which varies with angle (Day, 1973*a*) and its positive or negative contribution, depending on the orientation of the figure.

Determinants of illusion magnitude

Among the factors determining the size of the Poggendorff illusion which have been systematically investigated are the distance between parallels, the angle between transversals and parallels, and orientation of the whole figure.

As shown above the increasing illusion as the distance between the parallels increases (Burmester, 1896; Weintraub & Krantz, 1971; Hotopf & Ollerearnshaw, 1972*b*) results from the angular difference between actual and apparent oblique direction between the transversals. The basis of this is clear from the results represented in Fig. 5. The Poggendorff illusion also increases directly as a function of the obtuse angle between transversals and parallels or, alternatively, inversely with the acute angle between these two elements (Burmester, 1896; Velinsky, 1925; Weintraub & Krantz, 1971). This variation is attributable to both the OA and HV components. The OA effect increases with the size of the angle achieving a maximum at about 150°* and the illusion due to the HV effect alone is maximal when the transversals are aligned at 45°, i.e. 135° (Day, 1973*a*). Presumably the combination of the two functions determines the non-linear relationship between apparent misalignment and transversal–parallel angle, described by Velinsky (1925).

The Poggendorff illusion varies in size as the whole figure is rotated, achieving a maximum when the transversals are at about 45° and the parallels are vertical or horizontal and a minimum when the reverse obtains (Obonai, 1931; Leibowitz & Toffey, 1966; Weintraub & Krantz, 1971). The difference in the illusion when the parallels are vertical or horizontal compared with when they are oblique is marked and cannot be readily accounted for by presence or absence of the three components or by their individual variation. The HV component varies with orientation and, as pointed out above, can account for the relatively small difference in the illusion between the vertical and horizontal position of the parallels. However, the difference in size of the illusion between these and the oblique positions is considerably greater.

In proposing a basis for the orientation effect it is worth noting that a similar variation, least for the cardinal positions of the transversals and greatest for the oblique, occurs also for the Zöllner illusion. It is suggestive also that judgements of orientation (Jastrow, 1892; Gibson & Radner, 1937; Rochlin, 1955) and visual acuity (Hamblin & Winser, 1927; Higgins & Stultz, 1950; Leibowitz, 1955) are optimal when the stimulus elements correspond with the vertical and horizontal meridians of the eye and worst when they correspond with the 45° meridians.

The Poggendorff figure presents a visual acuity task, a point which has been entirely overlooked in earlier discussions. The subject is required to judge alignment as in a vernier acuity task. Vernier acuity varies with test pattern orientation with best acuity occurring when the figure is horizontal or vertical (Leibowitz, 1955). It is reasonable to suppose that in illusion figures in which the task involves orientation or direction judgements or visual resolution the illusion will be least when the judged elements correspond with the vertical or horizontal meridians and greatest when they correspond with the oblique meridians of the eye. Insofar as illusions derive from processes which distort visual space it can be expected that the processes exercise their greatest effect when acuity is poorest.

---

* In an unpublished experiment (Day, in preparation) 20 subjects adjusted the distance between a pair of dots to match the apparent extent of a 50 mm horizontal line with a 50 mm line joining it at angles between 15° and 180° in 15° steps. The greatest overestimation of extent occurred for the 150° condition.

DISCUSSION

The essential feature of this explanation is that the horizontal–vertical, longitudinal–transverse and obtuse angle effects are linear distortions of space which singly or jointly cause apparent relative elongation of a figure. As a result apparent oblique direction across the figure is modified so that two collinear elements appear misaligned. Three points which deserve further comment emerge from this argument: apparent direction in 'unchartered' space, the notion of basic illusion components, and relationships between illusions in terms of their basic components.

Numerous earlier explanations of the Poggendorff illusion rest on a change in the apparent orientation of the transversals and various causes of this change have been proposed. However, it is not the direction of the transversal which is modified (Day & Dickinson, 1975) but a direction, usually indicated by the transversal intersections, between the transversals. It seems, therefore, that misdirection in space occurs mainly in 'open' or 'unchartered' space in the absence of lines or edges which mark the true direction. That is, it is not a direction indicated by a continuous line which is modified by linear distortion in the figure but rather the direction marked by two terminal points, the points of departure and arrival, so to speak.

A clear implication of this explanation is that the HV, LT and OA effects are basic components and that the illusions of direction and alignment in certain figures are secondary outcomes. The question which immediately arises is whether some or all the classical geometrical illusions derive from one or more basic effects. In this regard it can be noted that there is now a body of convincing data to suggest that the classical Müller-Lyer illusion in the arrow-head figure involves at least two effects (Cooper & Runyon, 1970; Erlebacher & Sekuler, 1969, 1974; Sekuler & Erlebacher, 1971).

The possibility of relationship between visual illusions which appear on first consideration to be unrelated follows from the notion of basic components of illusions. For example, the Poggendorff misalignment effect and the Müller-Lyer size illusion have in common the OA component, although the effect of this component is different in the two cases. In one it results in an apparent misalignment and in the other a modification of apparent length. It seems possible that a basis for classifying illusions might emerge from a consideration of the contributing components rather than the more complex effects they generate. The components of geometrical illusions might be thought of as analogous to 'genotypes' and the more complex classical illusions as their 'phenotypical' outcomes. Thus in the Poggendorff figure the phenotypical effect of the genotypical OA component is misalignment and in the Müller-Lyer change in length. That is, the same component produces different effects according to the perceptual task which the observer undertakes.

REFERENCES

AVERY, G. C. (1969). The horizontal–vertical effects for judged extent and velocity. Unpublished PhD thesis, Monash University.

AVERY, G. C. & DAY, R. H. (1969). Basis of the horizontal–vertical illusion. *J. exp. Psychol.* **81**, 376–380.

BEGELMAN, D. A. & STEINFIELD, G. (1967). An investigation of several parameters of the horizontal–vertical illusion. *Percept. Psychophys.* **2**, 539–543.

Blakemore, C., Carpenter, R. H. S. & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature, Lond.* **228**, 37–39.

Burmester, E. (1896). Beitrag zur Aperimentellen Bestimmung geometrisch-optischen Täuschungen. *Z. Psychol.* **12**, 355–394.

Coooper, M. R. & Runyon, R. P. (1970). Error increase and decrease in minimal form of Mueller-Lyer illusion. *Percept. mot. Skills* **31**, 535–538.

Curthoys, I., Wenderoth, P. & Harris, J. (1975). The effects of the motion path and the length of the variable segment in the Poggendorff illusion without parallels. *Percept. Psychophys.* (in press).

Day, R. H. (1973*a*). The oblique line illusion: the Poggendorff effect without parallels. *Q. Jl exp. Psychol.* **25**, 535–541.

Day, R. H. (1973*b*). The Poggendorff illusion with obtuse and acute angles. *Percept. Psychophys.* **14**, 590–596.

Day, R. H. (1974). The Poggendorff illusion: apparent displacement of the oblique lines. *Aust. J. Psychol.* **26**, 49–59.

Day, R. H. & Dickinson, R. G. (1975). The Poggendorff illusion: misalignment which is not attributable to apparent orientation of the transversals. *Q. Jl exp. Psychol.* (in press).

Erlebacher, A. & Sekuler, R. (1969). Explanation of the Müller-Lyer illusion: confusion theory re-examined. *J. exp. Psychol.* **80**, 462–467.

Erlebacher, A. & Sekuler, R. (1974). Perceived length depends on exposure duration: straight lines and Müller-Lyer stimuli. *J. exp. Psychol.* **103**, 724–728.

Finger, F. W. & Spelt, D. K. (1947). The illustration of the horizontal–vertical illusion. *J. exp. Psychol.* **37**, 243–250.

Gibson, J. J. & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines: I. quantitative studies. *J. exp. Psychol.* **20**, 453–467.

Gillam, B. (1971). A depth processing theory of the Poggendorff illusion. *Percept. Psychophys.* **10**, 211–215.

Goldstein, M. B. & Weintraub, D. J. (1972) The parallel-less Poggendorff: virtual contours put the illusion down but not out. *Percept. Psychophys.* **11**, 353-355.

Hamblin, J. R. & Winsher, T. H. (1927). On the resolution of gratings by the astigmatic eye. *Trans. opt. Soc. Lond.* **29**, 28–42.

Helmholtz, H. (1886). *Handbuch de Physiologischen Optik*, 2nd ed. Hamburg and Leipzig: Voss.

Hering, E. (1861). *Beitrage zur Psychologie.* Leipzig: Engelman.

Higgins, G. C. & Stultz, K. (1950). Variations of visual-acuity with various test-object orientations and viewing conditions. *J. opt. Soc. Am.* **40**, 135–137.

Hotopf, W. H. N. & Ollerearnshaw, C. (1972*a*). The regression to right angles tendency and the Poggendorff illusion. I. *Br. J. Psychol.* **63**, 359–367.

Hotopf, W. H. N. & Ollerearnshaw, C. (1972*b*). The regression to right angles tendency and the Poggendorff illusion. II. *Br. J. Psychol.* **63**, 369–379.

Hotopf, W. H. N., Ollerearnshaw, C. & Brown, S. (1974). The regression to right angles tendency and the Poggendorff illusion. III. *Br. J. Psychol.* **65**, 213–231.

Imai, S. (1973). Experiments on Poggendorff illusion. *J. soc. Sci. Hum.* (Tokyo Metropolitan University), no. 90, 1–39.

Jastrow, J. (1891). A study of Zöllner's figures and other related illusions. *Am. J. Psychol.* **4**, 381–398.

Jastrow, J. (1892). On the judgment of angles and positions of lines. *Am. J. Psychol.* **5**, 214–221.

Judd, C. H. (1899). A study of geometrical illusions. *Psychol. Rev.* **6**, 241–261.

Krantz, D. H. & Weintraub, D. J. (1973). Factors affecting perceived orientation of the Poggendorff transversal. *Percept. Psychophys.* **14**, 511–517.

Künnapas, T. M. (1955). An analysis of the vertical–horizontal illusion. *J. exp. Psychol.* **49**, 134–140.

Láska, W. (1890). Über einige optische Ürteilstaüschungen. *Archs Physiol.* 326–328.

Leibowitz, H. (1955). Some factors influencing the variability of vernier adjustments. *Am. J. Psychol.* **68**, 266–273.

Leibowitz, H. & Toffey, S. (1966). The effects of rotation and tilt on the magnitude of the Poggendorff illusion. *Vision Res.* **6**, 101–103.

Müller-Lyer, F. C. (1889). Optische Ürteilstaüschungen. *Arch. Physiol.* Suppl. Band. 263–270.

OBONAI, T. (1931). Experimentelle Untersuchungen über den Aufbau des Schraumes. *Arch. ges. Psychol.* **82**, 308–328.

PIAGET, J. (1969). *The Mechanisms of Perception.* London: Routledge & Kegan Paul.

PIAGET, J. & DENIS-PRINZHORN, M. (1953). Recherches sur le développement des perceptions. XVI. L'ostination perceptive des côtés du rectangle. *Arch. Psychol.* **34**, 109–132.

PRESSEY, A. W. & DEWAR, M. J. (1970). A reversed Poggendorff illusion. *Psychon. Sci.* **20**, 227–228.

PRESSEY, A. W. & SWEENEY, O. (1972). Acute angles and the Poggendorff illusion. *Q. Jl exp. Psychol.* **24**, 169–174.

RESTLE, F. (1969). Illusions of bent line. *Percept. Psychophys.* **5**, 273–274.

ROBINSON, J. O. (1972). *The Psychology of Visual Illusion.* London: Hutchinson.

ROCHLIN, A. M. (1955). The effect of tilt on the visual perception of parallelism. *Am. J. Psychol.* **68**, 223–236.

SANFORD, E. C. (1898). *A Course in Experimental Psychology.* Boston: Heath.

SEKULER, R. & ERLEBACHER, A. (1971). The two illusions of Müller-Lyer: confusion theory re-examined. *Am. J. Psychol.* **84**, 477–486.

TOLANSKI, S. (1964). *Optical Illusions.* London: Pergamon.

TONG, L. & WEINTRAUB, D. J. (1974). Contour displacements and tracking errors: probing twixt Poggendorff parallels. *Percept. Psychophys.* **15**, 258–268.

VELINSKY, S. (1925). Explication physiologique de l'illusion de Poggendorff. *Ann. Psychol.* **26**, 107–116.

WEINTRAUB, D. J. & KRANTZ, D. H. (1971). The Poggendorff illusion: amputations, rotations, and other perturbations. *Percept. Psychophys.* **10**, 257–264.

WUNDT, W. (1893). *Grundzüge der physiologischen Psychologie*, 4th ed. Leipzig: Engelman.

(*Manuscript received* 3 *May* 1975; *revised version received* 31 *October* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 553–558
Printed in Great Britain

THE EFFECT OF TASK COMPLEXITY AND SEQUENCE ON RULE LEARNING AND PROBLEM SOLVING

By J. SWELLER
University of New South Wales

Two experiments were carried out in order to test the effects of task sequence on the speed of rule learning and problem solving. Expt. I involved numerical rule-learning tasks. One group of subjects was presented with the tasks using a simple–complex sequence while another group was presented with the complex rather than the simple task initially. The simple–complex order resulted in a far superior performance. Expt. II tested the effect of task difficulty and task precedence using problem-solving tasks. Speed of problem solving was not found to be significantly affected by task order.

These contrasting results were discussed in terms of hypothesis testing theory and learning hierarchies.

If a person is required to learn a series of related tasks which differ in terms of difficulty or complexity, the order in which the tasks are presented may affect the speed of learning. The total speed of learning the tasks may be more rapid if the easier tasks, for example, precede the more difficult. Gagné (1970) recommends this particular order in his theory of learning. As an example, simple learning processes such as discriminations should be learned before concepts, concepts should be learned before rules which involve the concepts and rules in turn should be learned before attempts are made to solve problems involving these rules. Furthermore, easier discriminations, concepts, rules and problems should be presented before, respectively, more difficult discriminations, concepts, rules and problems. The following evidence is available concerning this latter aspect of Gagné's theory.

Baker & Osgood (1954) and Restle (1955) both used discrimination transfer paradigms and found an easy–difficult sequence resulted in more rapid learning than the reverse sequence. (Baker & Osgood in fact compared an easy–difficult sequence against a group which was given the difficult task only.) House & Zeaman (1960) obtained similar results with retarded children on discrimination tasks. (All other studies reviewed here used normal adults.) Trabasso (1963) used a concept learning task and again easy–difficult superiority was obtained. Both of these experiments used a difficult task only control group. There appear to have been no studies carried out using a rule-learning task. Cook (1937) and Reed, Ernst & Banerji (1974) gave subjects problem-solving tasks. The results differed substantially from those summarized above. Cook used pyramid puzzles of varying complexity and found that the difficult–easy sequence was superior to the reverse in terms of errors, moves and time to solution. Unfortunately, these differences were small and were not analysed for significance. Reed *et al.* studied transfer from the Missionary–Cannibal problem to the Jealous Husbands problem and vice versa. These two problems are similar, with the Missionary–Cannibal problem being less complex. Results indicated no effect due to task order except for subjects who had been acquainted with the relation between the two tasks before being presented with the second problem. These subjects solved the problems more rapidly (in terms of seconds to solution) if they were given the more difficult problems first. There was no significant difference between these two groups in terms of total moves to solution.

In terms of Gagné's hierarchy of learning types, these results may be classified as follows. The recommended order of presentation for discrimination and concept-learning tasks is simple–complex. There is no evidence concerning rule-learning tasks. There is some evidence that it is advantageous for complex problem-solving tasks to precede simpler problems although it is possible that under most conditions the sequence is irrelevant. There is hence a clear difference between the simpler discrimination and concept-learning tasks on the one hand, and the more complex problem-solving tasks on the other, while there is no evidence concerning rule learning which, according to Gagné's hierarchy, is intermediate to concept learning and problem solving.

Predictions concerning the outcome of an appropriate rule-learning experiment may be derived from results obtained by Sweller (1973). Using a numerical rule-learning task, he found that subjects trained to a low criterion on initial tasks were unable to discover subsequent, related but more difficult rules. Subjects given training to a higher criterion on the same tasks readily discovered the more difficult rules. These results were interpreted in terms of Levine's (1971) hypothesis theory. It was suggested that subjects given a low criterion of learning before transferring to another task tend to reject the relevant hypothesis set, effectively making the subsequent tasks insoluble. Subjects given a higher criterion of learning, use previously learned rules to formulate hypotheses leading to the discovery of new related rules on new related tasks.

All subjects in the above experiment were given a simple–complex task sequence. Following this sequence should hence result in relatively rapid learning provided an appropriate criterion is used. The 'complex' tasks should in effect be made simple as a consequence of subjects being predisposed by the easy tasks to test relevant hypotheses. Their hypothesis sampling set should be reduced by previous success, but nevertheless still contain the correct hypothesis. Evidence for this reduction in the hypothesis set would be provided if the complex tasks are learned more rapidly than the previously learned simple tasks. If, alternatively, the complex tasks are given initially, the subjects should have difficulty finding the appropriate rule – a difficult rule-learning task being defined as one in which the hypotheses needed to discover the rule come from a very large or infinite hypothesis set.

The present experiment uses similar rule-learning tasks to the ones used by Sweller (1973). It is predicted that subjects given a simple–complex sequence will learn the tasks rapidly, due to a high degree of transfer from the simple to the complex tasks. This transfer should be due to a large reduction in subjects' hypothesis sampling set. On the other hand subjects given the complex task initially, should have difficulty learning the appropriate rule due to a very large hypothesis sampling set.

EXPERIMENT I

Method

Subjects

The subjects were 20 university students doing courses in Educational Psychology.

Materials

The stimulus materials consisted of 40 two-digit numbers. The numbers were chosen randomly except that digits 0 and 9 were excluded and no number could occur more than once. Each number was printed on a separate card and the cards ordered randomly.

Procedure

Subjects were shown the first card and told that another number could be calculated from the card number. They were then told the calculated number and informed that a rule connected the two numbers. For each subsequent card, they were required to respond with a number that they considered to be the calculated number. The experimenter then told them the correct number for that particular card. The next stimulus card was then exposed. If necessary, the experiment continued after the fortieth card by returning to the first card. Subjects could respond at their own rate.

There were three consecutive tasks and each subject continued on each task until he either attained a criterion of 10 consecutive correct responses, or until he had completed 40 trials on one of the tasks without attaining criterion. In the latter case, no additional tasks were given and the experiment was terminated for that subject. Subjects attaining criterion were immediately transferred to the next task until they had completed three tasks. The three rules were in order of complexity: (1) adding the first digit to the card number – if the card number was 43, the correct response was 47; (2) adding the first digit to the card number plus 1 – if the card number was 43, the correct response was 48; (3) adding the first digit to the card number plus 3.

Results

Of the ten Group SC subjects, eight solved the three tasks, with the mean trials to criterion on each of the tasks in order of presentation being 9·3, 2·9, and 2·7 respectively, giving a total mean for the eight subjects across the three tasks of 14·9 trials. The differences between the three means on the three tasks are significant, $F = 5{\cdot}86$, d.f. = 1, 7, $P < 0{\cdot}05$. An $X^1 = \log X$ transformation of the original data was carried out for purposes of this analysis in order to equalize variances. Conservative degrees of freedom are used since the covariance matrix was not tested for symmetry. Of the remaining two subjects in Group SC, one failed to solve the first task and the other the second task within 40 trials.

Of the ten Group CS subjects, two solved the three tasks and the remaining eight were unable to solve the first task they were given within 40 trials.

The difference between the two groups in terms of number of subjects able to solve all three tasks is significant using a Fischer exact probability test, $P = 0{\cdot}01$ (one-tailed).

Discussion

The results strongly indicate an easy–difficult sequence to be necessary for the type of rule-learning task used. Learning was rapid for all three tasks using this sequence, and transfer was sufficiently large to result in the more complex tasks being learned significantly more rapidly than the simple task. Of the 45 errors made by the eight subjects who completed the transfer tasks, 41 were caused by testing previously correct hypotheses. This supports the suggestion that subjects were sampling from a vastly reduced hypothesis set. In contrast, most subjects given a difficult–easy sequence were unable to discover the rule governing the most complex task – a task learned within a few trials by subjects required to learn easier rules first. In this case they were sampling from a large, possibly infinite hypothesis set.

Including the results of Expt. I with those reviewed above, there appears to be a distinction between discrimination, concept, and rule-learning tasks on the one hand, and problem-solving tasks on the other. Under these circumstances, the results of problem-solving tasks assume more than their normal quota of importance.

Expt. II is a partial replication of Cook's (1937) problem-solving experiment. There are several reasons for this replication. Firstly, as pointed out previously, Cook's results are ambiguous due to an inadequate data analysis. (Inadequate with respect to the problems considered here.) Secondly, before transferring his subjects to a more or less complex task, Cook required them to repetitively solve the same problem until it had been solved in the minimum number of moves possible. The same procedure was then followed on the new task. There was hence considerable positive transfer on each task as well as transfer between tasks. Since the latter is of primary interest, and since it may have been affected by the former, in the present experiment each subject solved both problems once only. Thirdly, there is some surface similarity between the pyramid puzzle and the rule-learning tasks of Expt. I. In both cases solution of the more complex task involves prior solution of the simpler task.

EXPERIMENT II

Method

Subjects

The subjects were 20 university students doing courses in Educational Psychology.

Materials

Five discs of graduated diameter were used. These could be placed in three circles. The circles formed a triangular configuration.

Procedure

Subjects were presented with four or five of the discs in one of the circles. The largest disc was on the bottom and the remaining discs were graduated in size to the smallest on top. Subjects were informed that their task was to move all of the discs so that they finished in the same order on either one of the other circles. They could only move one disc at a time and only if it had no discs above it. A larger disc could never be put on a smaller disc. Apart from these restrictions, any disc could be moved to any circle at any time. Subjects were informed that their number of moves and time to solution would be recorded.

Group 4–5 was given the four-disc problem first immediately followed by the five-disc problem. Group 5–4 was given the two problems in reverse order.

Results

The results in terms of mean time in seconds to solution are given in Table 1. A problem order (4- or 5-disc problem first) by problem position (initial or transfer problem) analysis of variance was carried out using an $X^1 = \log X$ transformation of the original data of each subject. This was done in order to equalize variances. The results of this analysis indicated no effect due to problem order $F < 1$; a significant effect due to problem position resulting in positive transfer, $F = 18{\cdot}9$, d.f. $= 1$,

Table 1. *Mean time (secs) to solution (mean moves to solution are in parentheses)*

	First problem presented	Second problem presented
Group 4–5	107 (19)	227 (45)
Group 5–4	374 (52·8)	42 (18·5)

18, $P < 0{\cdot}001$; a significant interaction effect due to differences in difficulty between the four- and five-disc problems, $F = 106{\cdot}9$, d.f. = 1, 18, $P < 0{\cdot}001$.

A similar analysis was carried out on the results of the number of moves to solution (Table 1). A logarithmic transformation was again used. The results indicated no effect due to problem order, $F < 1$; no transfer effect, $F = 1{\cdot}4$, d.f. = 1, 18, $P > 0{\cdot}05$; a significant interaction effect, $F = 283{\cdot}4$, d.f. = 1, 18, $P < 0{\cdot}001$.

Discussion

The results indicate task sequence to be largely irrelevant for the type of problem used. Total speed of learning both tasks was similar for both groups using either moves or time to solution as a dependent variable. These results probably accord with those obtained by Cook (1937) using a similar problem. (As stated previously, Cook did not analyse his results statistically with respect to the question at hand.) The results partly conform with those obtained by Reed *et al.* (1974). These investigators found no effect due to task precedence unless subjects were clearly acquainted with the relation between the two problems. In the latter case a complex–simple sequence was found to be superior.

The results also indicate that transfer from the first to the second task was primarily reflected in a decrease in the time to solution rather than a decrease in the number of moves to solution on the transfer task. This is primarily due to subjects' high level of performance on the initial task. The four-disc problem can be solved in a minimum of 15 moves and the five-disc problem in a minimum of 31 moves. As can be seen from Table 1, the mean moves to solution on Problem 1 do not permit considerable improvement on Problem 2. Presumably subjects were willing to sacrifice time in order to reduce their number of moves on the first problem. This would result in improvement on Problem 2 being in terms of time rather than moves.

While sequence did not differentially affect total speed of learning, it did affect degree of transfer, using time to solution as a measure of problem-solving efficiency. The percentage of transfer for Group 4–5 was 24 per cent, while that for Group 5–4 was 44 per cent where percentage of transfer is calculated using Murdock's (1957) formula:

$$\frac{\text{control} - \text{experimental}}{\text{control} + \text{experimental}} \times 100.$$

The initial problem for each group acts as a control for the transfer problem of the alternate group. There was hence substantially more transfer when the difficult problem was given initially. A similar calculation is not carried out using moves to solution due to the non-significant transfer effect using this measure.

GENERAL DISCUSSION

The results of Expts I and II contrast. Subjects tended not to be able to discover a higher-order rule which incorporated a lower-order rule unless they first learned the latter. On the other hand, solving a lower-order problem before a higher-order problem which incorporated the former, conferred no more advantage to subjects than the reverse sequence. The results of the rule-learning experiment (Expt. I) conform with previous experiments using discrimination and concept-learning tasks, and with Gagné's (1970) theory of learning, while those of the problem-solving experiment (Expt. II) do not. This distinction can be satisfactorily explained by assuming that subjects faced with discrimination, concept- or rule-learning tasks, test clear-cut hypotheses on each trial. As stated previously, hypothesis testing should lead to more rapid learning if a simple–complex sequence is used. Problem solving, of the type used in Expt. II, does not involve the testing of a variety of clear-cut hypotheses. No single, relatively simple hypothesis which will result in a solution to the problem is used. Since hypothesis testing does not occur, task sequence becomes irrelevant to total speed of learning.

REFERENCES

BAKER, R. A. & OSGOOD, S. W. (1954). Discrimination transfer along a pitch continuum. *J. exp. Psychol.* **48**, 241–246.

COOK, T. W. (1937). Amount of material and difficulty of problem solving. *J. exp. Psychol.* **20**, 288–296.

GAGNÉ, R. M. (1970). *The Conditions of Learning.* London: Holt, Rinehart & Winston.

HOUSE, B. J. & ZEAMAN, D. (1970). Transfer of a discrimination from objects to patterns. *J. exp. Psychol.* **59**, 298–302.

LEVINE, M. (1971). Hypothesis theory and non-learning despite ideal S–R-reinforcement contingencies. *Psychol. Rev.* **78**, 130–140.

MURDOCK, B. B. (1957). Transfer designs and formulas. *Psychol. Bull.* **54**, 313–326.

REED, S. K., ERNEST, G. W. & BANERJI, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychol.* **6**, 436–450.

RESTLE, F. (1955). A theory of discrimination learning. *Psychol. Rev.* **62**, 11–19.

SWELLER, J. (1973). Effect of amount of initial training on concept shift problems. *J. exp. Psychol.* **99**, 134–136.

TRABASSO, T. R. (1963). Stimulus emphasis and all-or-none learning in concept identification. *J. exp. Psychol.* **65**, 398–406.

(*Manuscript received* 5 *May* 1975; *revised manuscript received* 21 *October* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 559–567
Printed in Great Britain

RESPONSES TO HOMOPHONES AS A FUNCTION OF CUE WORDS ON THE UNATTENDED CHANNEL

BY S. H. A. HENLEY
Department of Psychology, University College, London

A prediction based on the model of attention advanced by Dixon (1971) was tested: namely, that responses to homophones presented to one ear, at supraliminal intensities, would be influenced by subliminal cue words presented to the other ear. Considerable support was found for the hypothesis in terms of response latencies, but not in terms of verbal content. It is suggested that these data make it possible to reconcile apparent discrepancies between the results of other recent studies of dichotic listening.

The view of those attention theorists (Deutsch & Deutsch, 1963; Norman, 1968; Dixon, 1971) who maintain that selection among inputs follows a preliminary semantic analysis of *all* signals, both attended and unattended, has received support from recent studies of dichotic listening (Lewis, 1970; Corteen & Wood, 1972; Mackay, 1973; Henley & Dixon, 1974), and dichoptic viewing (Somekh & Wilding, 1973). Data from these studies have suggested not only that material, of which subjects remain unaware, on an unattended channel, is analysed for meaning, but moreover that it may be integrated with primary material on the attended channel when it is relevant to the ongoing task. Mackay, for example, has shown that a word on the unattended channel in dichotic listening will serve to reduce the ambiguity of an attended sentence.

According to the model of attention put forward by Dixon (1971), a supraliminal stimulus, having passed through analysers to dictionary units, ultimately activates conceptually associated units while at the same time setting in motion a central control mechanism which, acting upon information from various inputs (relating to past and ongoing events, motivational states, etc.), selectively inhibits such conceptual associates as are not relevant to the task in hand. A subliminal stimulus, on the other hand, activates the conceptual associates, but is of insufficient strength to set the central control mechanism in operation. Under these conditions, sensory data are analysed and classified, but nothing is blocked.

The present experiment tested a prediction from this theory that responses to homophones presented to one ear, at supraliminal intensities, should be affected by subliminal cue words on the other ear. It was postulated that a supraliminal homophone would activate associates belonging to two spheres of meaning. Assuming responses to either meaning of the homophone to be equiprobable, there would be no *a priori* reason why any one set of conceptual associates should be inhibited more than the other. Under these circumstances, the ambiguity should be resolved by a subliminal cue word activating *its* associates. The latter would, presumably, overlap one set of those associates activated by the supraliminal stimulus, and as a result of the 'extra strength' thus gained, it would be to this meaning of the homophone that the subject should respond. By the same token, *if* the subliminal stimulus *did* reduce the ambiguity facing the central control at the level of the conceptual associates, response latencies should be shorter if the cue was used to facilitate performance than if it was not. It was further predicted that these effects would be most apparent if the subliminal stimulus preceded the homophone by a short time interval.

METHOD

Materials

On the basis of pilot studies, a list of 14 homophones was prepared, together with two lists of matching cue words whose meanings will henceforth be labelled 'A' and 'B'. The homophones and their respective cues are given below:

Homophone	Cue (meaning 'A')	Cue (meaning 'B')
1. WEEK/WEAK	DAY	FEEBLE
2. BEAR/BARE	ANIMAL	NAKED
3. PEAR/PAIR	FRUIT	TWO
4. DAM/DAMN	WATER	SWEAR
5. PALE/PAIL	WHITE	BUCKET
6. PEACE/PIECE	WAR	PORTION
7. WEIGHT/WAIT	HEAVY	STAY
8. HOLE/WHOLE	CAVITY	COMPLETE
9. SEA/SEE	WATER	LOOK
10. SUN/SON	MOON	DAUGHTER
11. MAID/MADE	SERVANT	PRODUCED
12. SAIL/SALE	BOAT	SHOP
13. DEER/DEAR	ANIMAL	DARLING
14. MAIN/MANE	CHIEF	LION

Apparatus

A Revox tape recorder (stereo model A77) was used for presentation of the material. The two lists of cue words were recorded four times each on channel 1, sound level being kept constant by means of an automatic gain control. List 'A' was recorded first in one order, then in the reverse order; list 'B' was then recorded likewise; finally, the whole sequence was repeated. The interval between each word was 10 sec. The list of homophones was subsequently recorded on channel 2 (a total of eight times), at a similar level to the cue words. The first four lists were recorded such that each homophone on channel 2 was synchronized with the appropriate cue word on channel 1 ('simultaneous' condition); each homophone in the remaining four lists was timed to occur 5 sec *after* its cue word on channel 1 ('delayed' condition). Five-second bursts of white noise were also recorded on channel 2, between the homophones, to prevent 'cross-talk' in the 'delayed' condition.

The output from channel 1 of the Revox was led, via a suitable matching attenuator, to a Peters audiometer, which served as a calibrated attenuator, and thence to one headphone. Channel 2 was led, via a second matching attenuator, to the other headphone.

A second output from channel 2 was led to a Philips stereo tape recorder which was also used to record subjects' responses, and from which reaction times were taken subsequently.

Subjects

The 16 men and 12 women who acted as subjects included two members of staff from the psychology department, and students from various other departments of University College London, three volunteer nurses from University College Hospital, and a speech therapist. Three subjects were rejected from the final analysis on the grounds of auditory defects which prevented them from responding to almost 50 per cent of the homophones. Of the remaining 25 subjects, ten comprised a control group, and the other 15 were randomly assigned to one of two experimental groups, A and B, according to which set of cue words were on the unattended channel. These two groups were further subdivided in terms of the condition under which they were to perform – 'simultaneous' or 'delayed'. All subjects were paid for their participation.

Procedure

On arrival at the experimental room, subjects were asked to recline on a portable bed, and given the following instructions:

'This experiment is concerned with subjects' ***immediate*** responses to a fairly short list of very common words. All you have to do is to lie back, relax, and say the first thing that comes into your head each time you are presented with a word. The words will be presented through headphones at the rate of one every ten seconds; each one will be preceded by a burst of white noise which serves as a warning signal. Before we begin, however, it's necessary to get some measure of your auditory sensitivity to ensure that you won't be deafened by the white noise. When I first switch on the tape, you will hear a list of numbers. All you have to do is to turn this dial (subject shown dial on attenuator) in a clockwise direction until you are certain you can no longer hear them. When you are sure you have reached this point, turn the dial for one more click so that we can get the lowest baseline measure of your sensitivity. (The numbers were recorded on channel 1 at the same level as the subsequent cue words. The last part of the procedure ensured that the latter were at least 5 dB below each individual's awareness threshold.) After this, there will be a short delay before the first word is presented. Just lie back and relax.'

Help was then given to each subject with putting on the headphones. In this way it was possible to ensure that half the subjects heard the homophones through the right ear, and half through the left ear (to counterbalance for a right ear advantage). The room was subsequently darkened and the tape started. As well as being recorded on the subsidiary tape recorder, each subject's responses were also noted by the experimenter on a record sheet which was used during the next stage of the experiment.

Following the first part of the procedure, which lasted approximately 3–4 min, subjects were told: 'I'm going to read each of your responses back to you, and together with each one, present you with two additional words. I would like you to say to which of these two words you think your response is most closely associated. For example, if your response was "chair", and I gave you as alternatives "table" or "sky", I imagine you would match your response to "table".'

The subject was subsequently given his own response to each homophone along with the two appropriate cue words. The order in which these cues were presented was reversed for alternate subjects. On the rare occasions when a subject found the choice difficult, he was asked to make a guess. This matching behaviour constituted one measure of the meaning to which each subject had been responding.

Finally, each subject was asked to write down as many as he could remember of the words presented in the initial task. He was informed that 'This is not a test of recall as such: it is simply for the purpose of providing an indication of which of the words most people are spontaneously able to remember'. When the subject claimed to be unable to remember any more, he was asked to draw a line under those recalled, and write down any omitted from the total list which was then read out to him. The spelling of these written responses constituted a second measure of the meaning to which each subject had been responding.

The procedure was identical for all subjects except that those in the control group received no input from channel 1 of the Revox (cue words), and experimental subjects were subdivided both in terms of the meaning of the cue words fed in on this channel ('A' or 'B'), and again in terms of the condition under which they performed ('simultaneous' or 'delayed').

RESULTS

Word association data

That the subliminal stimuli had not been effective in influencing verbal responses in the manner predicted was suggested by the fact that no significant difference obtained between the proportion of responses matched with meaning 'A' cues, and the proportion of responses matched with meaning 'B' cues by subjects in the control and experimental groups as a function of the subliminal input. Nor was the effect of the latter found to be any more apparent in the responses of experimental subjects who performed under the 'delayed' condition than in the responses of those who performed under the 'simultaneous' condition.

The word association data of the control group, however, suggested that a highly significant response bias was operating towards giving meaning 'A' responses. All ten of the control subjects gave more 'A' than 'B' responses, mean proportion of responses matched against meaning 'A' cues being 0·73, and mean proportion of responses matched against meaning 'B' cues being 0·27. This difference is significant at the 0·2 per cent level of confidence (two-tailed) by a sign test.

This bias was totally unexpected in view of the fact that the two meanings of the homophones were matched in terms of word frequency. However, a *post hoc* examination of the cue words indicated that list 'B' contained more cues which were synonyms of the homophones than list 'A', which contained a greater number of cues standing in a superordinate–subordinate relationship to the homophones. It is likely that associations to a word with its synonym are infrequent. This was evidently a possibility which should have been taken into account when the lists of cue words were prepared.

Written response data

The second measure of the meaning of the homophones to which subjects had been responding, namely the spelling of these words in the written 'recall' task, afforded an opportunity of examining whether the effects of the subliminal cue words upon voluntary response behaviour were greater after a delay of the order of minutes. Such a finding would be consistent with the finding by Silverman *et al.* (see Silverman, 1971) that the effects of subliminal stimuli are frequently delayed. It might also explain the failure of the subliminal cues to have a significant influence upon performance in the word association task.

Table 1 shows the mean proportion of responses with meaning 'A' and meaning 'B' spellings given by experimental and control subjects. (Data from this task are available for only seven of the control subjects: the remaining three did not complete this part of the procedure.)

The difference in the distribution of 'A' and 'B' responses between the two experimental groups, A and B, narrowly missed significance ($U = 13{\cdot}5$, $P < 0{\cdot}06$, one-tailed); nor did the mean proportion of responses with 'A' and 'B' spellings given by group A differ significantly from the mean given by the control group. The results of group B on this task, however, differed from those of the control group to a degree which was highly significant. Thus, subjects in group B gave a significantly *higher* proportion of responses with meaning 'B' spellings, and a correspondingly *lower* proportion of responses with meaning 'A' spellings than control subjects ($U = 4$, $P = 0{\cdot}003$, one-tailed). It seemed reasonable to conclude from this result that the subliminal cues *had* been effective in influencing performance on this task, at least among subjects in group B.

As in the word association task, the effects of the subliminal cues were no more apparent in the data of experimental subjects who performed under the 'delayed' condition than in the data of those who performed under the 'simultaneous' condition.

Table 1. *Mean proportion of responses with 'A' and 'B' spellings given by experimental and control subjects in written 'recall' task*

Control group (*n* = 7)		Experimental groups			
		Group A (*n* = 8)		Group B (*n* = 7)	
Homophone to Ear 1		Homophone to Ear 1		Homophone to Ear 1	
——		Cue 'A' to Ear 2		Cue 'B' to Ear 2	
'A'	'B'	'A'	'B'	'A'	'B'
0·79	0·21	0·73	0·27	0·62	0·38

Table 2. *Overall mean RTs (sec) for 'hits' and 'misses' given by experimental subjects*

	'Hits' (Responses matched against cue presented)	'Misses' (Responses matched against alternative cue)
RT (*n* = 15)	1·81	2·19

Reaction time data

Individual differences in response times precluded a direct comparison of RTs across experimental and control groups. In the attempt to overcome this problem, the data were dealt with in two ways: first, on the assumption that if the subliminal cues were utilized, RTs for 'hits' (i.e. 'A' responses by subjects in group A, 'B' responses by subjects in group B) should be faster than RTs for 'misses' ('B' responses by group A, 'A' responses by group B), each subject's mean RT for 'hits' was compared with his mean RT for 'misses'. As a second check on the effects of the subliminal stimuli upon RTs, the responses given to *each homophone* by the 15 experimental subjects were subdivided into 'hits' and 'misses', and the mean RT for each category calculated. These two means (for each word) were compared both with each other, and with the mean RT given by subjects in the control group for the same homophone. Thus, by these two methods, it was possible to examine the effects of the subliminal cues upon RTs both within and across subjects.

Table 2 shows the overall mean RTs given by experimental subjects for 'hits' and 'misses'.

Fourteen of the 15 experimental subjects gave faster mean RTs for 'hits' than for 'misses'. A Wilcoxon signed-ranks matched-pairs test showed that the difference was highly significant ($T = 9$, $Z = 2{\cdot}89$, $P = 0{\cdot}0019$, one-tailed).

No significant differences obtained between the RT data of those subjects who performed under the 'delayed' condition and those who performed under the 'simultaneous' condition.

Table 3 shows the mean RT for responses matched against the cue presented ('hits' across experimental subjects), the mean RT for responses matched against the alternative cue ('misses' across experimental subjects), and the mean RT for responses given under the control condition for each homophone. (Since the overall

Table 3. *Mean RT (sec) for responses correctly matched against cue presented, responsse matched against alternative cue, and responses given under the control condition*

	Experimental conditions		
Homophone	'Hits' (responses matched against cue presented)	'Misses' (responses matched against alternative cue)	Control condition
WEEK/WEAK	1·6	1·8	1·83
BEAR/BARE	1·72	2·44	2·15
PEAR/PAIR	1 68	2·6	1·84
DAM/DAMN	2·03	2·0	2·66
PALE/PAIL	1·74	2·11	1·36
PEACE/PIECE	1·85	2·1	2·31
WEIGHT/WAIT	1·74	2·13	2·2
HOLE/WHOLE	1·71	2·45	1·79
SEA/SEE	1·4	1·72	1·99
SUN/SON	1·71	1·9	1·91
MAID/MADE	2·32	2·5	2·07
SAIL/SALE	1·65	1·92	1·76
DEER/DEAR	1·37	2·86	1·71
MAIN/MANE	2·2	2·03	1·8
	24·72	30·56	27 38
$\bar{X}$	1·76	2·18	1·95

mean RT of the control subjects exceeded that of the experimental subjects by 6·2 per cent, the mean RT given by the control group for each homophone has been corrected by this amount in order to make it more directly comparable with the corresponding values for 'hits' and 'misses' given by the experimental subjects.)

It can be seen from this table that, for 12 of the 14 homophones, the mean RTs for responses matched against the cue presented were faster than the mean RTs for responses matched against the alternative cue. The difference was highly significant (Wilcoxon test, $T = 3$, $Z = 3{\cdot}11$, $P = 0{\cdot}0009$, one-tailed). Since the mean RTs of subjects in group A, over all words, did not differ from those of subjects in group B, the faster mean RTs for 'hits' than for 'misses' could not be attributed to the fact that group A produced more 'hits' than group B.

The mean RT for responses matched against the cue presented were also faster than the mean RTs for responses given under the control condition to 11 of the 14 homophones. This difference was also significant ($T = 23$, $Z = 1{\cdot}83$, $P = 0{\cdot}03$, one-tailed). Mean RTs for responses matched against the alternative cue word tended to be longer than those given under the control condition. This difference did not reach significance, however.

Introspective data

All subjects firmly denied awareness of any input on the unattended channel. As in a previous experiment involving the same paradigm as that employed in the present study (Henley & Dixon, 1974), several subjects went so far as to suggest that the experimenter had failed to handle the equipment properly, assuming this to be the reason for input not being in stereo. Only one subject (a psychologist) commented that if she was forced to make a guess she would report 'signal present'.

This subject was, in fact, one of the control group, and had received no input other than the supraliminal homophones! This fact, confirming as it does Dixon's (1971) observation that if someone, especially an authority figure, says 'Did you see (or hear) that?', 'the tendency is to respond in the affirmative, whether or not one has actually had the requisite experience' (p. 47), suggests that the remaining subjects' denials of awareness of material on the unattended channel can be considered a valid indication that the cue words on this channel were indeed subliminal.

DISCUSSION

The present findings appear fully consistent with the results of previous studies of a similar nature to the one reported here. Thus, on the one hand, bearing in mind the pre-existing response bias manifested by the control group, the failure of the subliminal cues to have an observable effect upon verbal behaviour accords with numerous demonstrations (e.g. Smith, Spence & Klein, 1959; Goldstein & Barthol, 1960; Goldstein & Davis, 1961) of the fact that 'in competition with stronger or incompatible tendencies, the effects of marginal stimuli are minimal or non-existent' (Dixon, 1971). The written response and reaction time data, on the other hand, suggest that the subliminal cues *were* analysed for meaning: in so doing they, like the results from other recent researches on dichotic listening and dichoptic viewing (Lewis, 1970; Corteen & Wood, 1972; Mackay, 1973; Somekh & Wilding, 1973; Henley & Dixon, 1974), appear to support those theories of attention (Deutsch & Deutsch, 1963; Norman, 1968; Dixon, 1971) which maintain that *all* inputs undergo a preliminary semantic analysis, as opposed to an analysis of only their purely physical characteristics, prior to selection.

An explanation of the present data in terms of responses to partial (physical) cues is rendered difficult on two accounts: first, the subliminal cue words on the unattended channel were fed in *at least* 5 dB below the intensity at which the subject was able to report awareness of *anything* on that channel; secondly, if subjects were responding to partial cues, one might expect that the stimulus words which gave rise to responses which were 'hits' would be the same for all subjects within each experimental subgroup: this was not the case.

Ideally, to increase reliability, the procedure used for determining threshold should have been repeated at the end of the experiment. Although this was not done, the subjects were questioned closely as to whether they had been aware of anything other than the supraliminal homophones at any time during the experiment. Since (as noted above) the tendency is to respond in the affirmative when confronted with the question 'Did you see/hear something?', and since there would seem little reason to doubt the honesty of the subjects on this score, their firm denials of awareness of anything at any time on the unattended channel can be fairly safely assumed to reflect the fact that the subliminal cues remained below the awareness threshold throughout the course of the experiment.

The fact that, over all homophones, mean reaction times for responses subsequently matched against the cue presented were faster than those for responses incorrectly matched, *and* for responses given under the control condition, supports the view that not only were the subliminal cues analysed for meaning, but that their utilization had a facilitating effect upon responding over and above any effects due to a response bias. The additional facts (*a*) that the mean RTs for responses incorrectly matched (i.e. ones which, presumably, had not utilized the cues) tended to be longer than the mean RTs for responses given under the control condition, and (*b*) that all but one of the experimental subjects gave faster mean RTs for 'hits' than for 'misses', militate against the possibility that reduced response latencies were simply a function of the fact that the brain had recorded the presence of a signal. Material on the unattended channel must have undergone an analysis for meaning.

Since Dixon's (1971) model of attention, unlike any of the others mentioned, is explicit with respect to the analysis of inputs which fail to achieve phenomenal representation, it seems particularly adequate to deal with the data reported here.

That performance on the written 'recall' task was influenced by the subliminal cues, in the direction predicted, to a greater extent than performance on the word association task, suggests that the more specific effects of the subthreshold stimuli (at least upon voluntary responses) were *delayed*. This conclusion receives some support from the results of researches by Silverman *et al.* (see Silverman, 1971), which have shown that when the effects of drive-related subliminal stimuli are sought on more than one task, they are more apparent in the *later* tasks given.

Finally, it is clear from the present results that reaction times constitute a very much more sensitive indicator of the effects of subliminal (or unattended) stimuli than do verbal responses. This finding, which is consistent with those of Spence (1961) and Spence & Bressler (1962), suggests that discrepancies between the results of different experiments on subliminal perception may arise in part from the use of different response indicators. It also makes it possible to reconcile the apparently disparate elements in the results of the dichotic listening experiments by Corteen & Wood and Lewis. While the former researchers reported that the shadowing of prose on the attended channel remained unaffected by shock-associated words on the unattended channel (and this *despite* the fact that the latter words, *and* their associates, continued to evoke significant G.S.R.s), Lewis's data suggested that words presented on the unattended channel can *interfere* with the shadowing of attended words. In the light of the present findings, it seems likely that Corteen & Wood might have found similar evidence of interference had they, like Lewis, used reaction times as a measure of the effects of unattended stimuli.

I am indebted to Dr N. F. Dixon for his help and encouragement during the preparation of this report, to Dr A. R. Jonckheere for advice concerning statistics, and to the Medical Research Council for financial support while the experiment was being carried out.

REFERENCES

CORTEEN, R. S. & WOOD, B. (1972). Autonomic responses to shock-associated words in an unattended channel. *J. exp. Psychol.* **94**, 308–313.

DEUTSCH, J. A. & DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.* **70**, 80–90.

DIXON, N. F. (1971). *Subliminal Perception: the Nature of a Controversy.* London: McGraw-Hill.

GOLDSTEIN, M. J. & BARTHOL, R. P. (1960). Fantasy responses to subliminal stimuli. *J. abnorm. soc. Psychol.* **60** (1), 22–26.

GOLDSTEIN, M. J. & DAVIS, D. (1961). The impact of stimuli registering outside of awareness upon personal preferences. *J. Personality* **29**, 247–257.

HENLEY, S. H. A. & DIXON, N. F. (1974). Laterality differences in the effect of incidental stimuli upon evoked imagery. *Br. J. Psychol.* **65** (4), 529–536.

LEWIS, J. L. (1970). Semantic processing of unattended messages using dichotic listening. *J. exp. Psychol.* **85** (2), 225–228.

MACKAY, D. G. (1973). Aspects of the theory of comprehension, memory, and attention. *Q. Jl exp. Psychol.* **25** (1), 22–40.

NORMAN, D. A. (1968). Toward a theory of memory and attention. *Psychol. Rev.* **75**, 522–536.

SILVERMAN, L. H. (1971). An experimental technique for the study of unconscious conflict. *Br. J. med. Psychol.* **44**, 17–25.

SMITH, G. J. W., SPENCE, D. P. & KLEIN, G. S. (1959). Subliminal effects of verbal stimuli. *J. abnorm. soc. Psychol.* **59** (2), 167–176.

SOMEKH, D. E. & WILDING, J. M. (1973). Perception without awareness in a dichoptic viewing situation. *Br. J. Psychol.* **64** (3), 339–349.

SPENCE, D. P. (1961). The multiple effects of subliminal stimuli. *J. Personality* **29**, 40–53.

SPENCE, D. P. & BRESSLER, J. (1962). Subliminal activation of conceptual associates. *J. Personality* **30**, 89–105.

(*Manuscript received* 4 *April* 1974; *revised manuscript received* 27 *September* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 569–577
Printed in Great Britain

THE RELATIVE DIFFICULTY POSITION FUNCTION

BY VICTOR M. AGRUSO, JR.
Drury College, Missouri
AND MARK D. RECKASE
University of Missouri, Columbia

A study was conducted to determine if the serial position curve remains invariant across developmental levels under different rote verbal learning tasks when the curves are expressed as Indices of Relative Difficulty. Eighty students, ages 5–11, from the University of Missouri-Columbia Laboratory School served as subjects for this study. Each subject participated in two experimental tasks, multitrial free recall and serial recall, counterbalanced in presentation. Seven words for each of the two experimental tasks and for each of the two practice tasks constituted the list length. On each task, every subject received a different random order of the list. The data consisted of an error score, cumulated over five trials, at each of seven serial positions for each subject in two tasks. The data were transformed into the Index of Relative Difficulty, which was the method used in representing the serial position effect. The statistical analyses were a multiple discriminant analysis on the four age groups for each of the two tasks and a Hotelling T^2 analysis on the serial versus free recall mean vectors. All analyses were performed on the Index of Relative Difficulty. The analyses indicated that age or grade level in and of itself is not a significant factor in changing the bowed skewed curve found in rote verbal learning tasks. The effect of serial position on the shape of the serial position curve has been shown to be consistent with the findings of other investigators.

McGeoch & Irion (1952) report that since the early investigation by Ebbinghaus in the 1870s, researchers have found that when a subject memorizes a list of nonsense or meaningful materials he learns the beginning (primacy) and the end (recency) more readily than the middle position. If the absolute numbers of errors made per position are plotted, a serial position curve invariably results, i.e. a curve showing bowness and skewness with beginning items having the least errors, middle items the most, and those at the end having the next least errors.

The relative invariance of position effects (on serial position curves) in rote verbal learning has been demonstrated under numerous conditions utilizing a wide variety of stimulus materials (Hovland, 1938; McCrary & Hunter, 1953; Braun & Heymann, 1958; Jensen, 1962*a*). Various theories have been advanced to account for the 'serial position effect'; however, until 1952 the Lepley–Hull theory had received the most verification or at least had been the most tenable. This theory (Hull, 1935) essentially posits that the serial position curve is a function of the amount of inhibition associated with the responses in the middle of the list, i.e. intralist interference. The many experiments conducted to test this and various other hypotheses yield results that indicate the shape of the curve is invariant under a wide variety of procedures, such as massed and distributed practice, instructions, meaningfulness of items, retention interval, interitem interval, position, length of list, differences in learning ability, and extra-experimental interference beyond formal retroactive and proactive inhibition.

Jensen (1962*b*) has seriously questioned whether the serial position curve is indeed invariant. He suggests that although the McCrary–Hunter (1953) and the Braun–Heymann (1958) studies were clever representations of a new look at an old problem, the results were none the less based on fallacious analyses. In those studies the curves were plotted not in terms of mean number of absolute number of errors made at each position, but rather in terms of the percentage of total errors made at each serial position. Thus, under this type of analysis the shapes of the curves for learning under varied conditions and with varied stimulus materials were essentially identical.

Jensen (1962*b*, *c*) suggests that the curve can also be demonstrated in terms of the number (or percentage) of correct responses at each serial position (also reported by McGeoch & Irion 1952, p. 116). He further suggests that the error plot and correct responses plot (regardless of the method of expression) for the same set of data are not always the identical shape, i.e. the bowness might differ from one curve to another. Thus, a weakness in those studies is that the invariancy of the curve has not been shown in terms of a correct response plot.

Murdock (1960, p. 24), in reaction to the seemingly paradoxical experimental findings with respect to the serial position curve, stated '. . . it would almost seem that the shape of the serial position curve has nothing whatsoever to do with learning . . . since it seems to be unaffected by variables that are known to affect rate of learning'. Murdock suggests that the curve is a function of the unequal distinctiveness of the items due to their ordinal position within the list, i.e. initial items are more distinctive than the middle items. Thus it would seem appropriate to go beyond the learning variables. If the curve does not seen to be a function of learning, perhaps it may be a function of developmental level.

Spiker (1960) and Goulet (1968) point out that while the literature on serial verbal learning in adults is abundant, the research with children is sparse. Spiker suggests that this situation could be partially due to the fact that the reading ability of elementary school children usually has been considered inadequate for the efficient study of verbal learning. Furthermore, since the serial task is a relatively unanalytical tool (each item having the double function of stimulus and response), it is less desirable than paired associates (PA) learning. There is a consensus among most investigators that children, like adults, produce a bowed serial position curve (Andrews, Shapiro & Cofer, 1954; Barnett, Ellis & Pryer, 1960). Lepley (1934, p. 46), however, has observed changes in the shape of the curve as a function of age. In one of his experiments a group of seventh grade subjects exhibited the primacy effect to a lesser degree than a group of eleventh grade subjects. This evidence suggests that serial learning order varies with the learner's age. He states that the gross factor of age must be taken into account in future study of serial learning order.

A study of the relationship between age and the serial position effect would add to our knowledge about the invariance of the curve as it relates to age. Furthermore, the serial position effect has important implications for the conditions under which learning occurs, since the learning of verbal chains (sequences) represents an integral part of the educational process.

In summary, the experimental findings demonstrate that the skewness and bowness of the curve is found regardless of the nature of stimulus materials (nonsense syllables, proper names, words, objects, figures) while the degree of bowing was shown to vary as a function of intertrial and interitem interval, distribution of practice, item familiarity, learning ability. One of the critical variables seems to be method of expressing the experimental results.

The purpose of this study was to determine whether or not the shape of the serial position curve is an invariant function of rote verbal learning across varying developmental levels under different presentation modes when the curves are expressed as Indices of Relative Difficulty. Research to this time indicates that the shape of the serial position curve is a relatively invariant function across a wide variety of conditions. However, the data come primarily from investigations using adult populations.

METHOD

Subjects

Eighty subjects from the University of Missouri-Columbia Laboratory School participated in this study. There were 20 subjects at each of the following developmental (age or grade) levels: kindergarten (K1 which in effect is pre-school since the age of the K2 group is 6 while for K1 the age is 5); grade 1; grade 3; and grade 5. The grade levels correspond approximately to ages 5, 7, 9, and 11 respectively. Control groups were not required because each subject served as his own control, receiving both the serial and free recall tasks.

Design

Through the procedure given by Jensen (1962*b*), the data were transformed into the Index of Relative Difficulty. The procedure for obtaining the Index for a single subject is:

(*a*) Correction factor. Correction factor $= 1/P(0{\cdot}50R-E)$, where P = number of positions, R = total number of responses, and E = number of errors.

(*b*) Algebraically add the correction factor to the number of errors at each position.

(*c*) After performing (*b*), obtain the sum total over all positions.

(*d*) Divide (*c*) into each position, (*b*), to convert the quantities into percentages which will sum to 100 per cent.

(*e*) The group curve is obtained by averaging (*d*) for each position over all subjects.

'The Index has the advantage that it shows the "shape" of the serial curve independently of difficulty (number of errors) and efficiency (percentage of errors). Also, the curves based on errors or correct responses are exactly the same shape. The method permits direct comparisons of serial curves obtained under any conditions. In order that the serial curve of each subject is equally weighted in the average curve for the group, the Index is applied to the data for each subject. The points on the curve are then converted to percentages for each subject and are averaged over all subjects to produce the group curve, which is an Index of Relative Difficulty.'

Apparatus

The word lists were presented auditorally to each subject using a tape recording. A Norelco Tape Cassette, Model El 3302 was utilized both in recording and playing back the lists which were equally spaced with respect to time.

Procedure

In the initial phase of the study word lists were constructed for both the practice and experimental tasks. It was necessary to construct the lists so that they would not be too easy for the older subjects and, at the same time, not too difficult for the younger subjects. The words were taken from the Thorndike & Lorge (1944) pool. Seven words for each of the two experimental tasks and for each of the two practice tasks constituted the list length. All words have a G count of AA and are one syllable nouns. The words were also selected so that no repetition of initial letters and no meaningful adjacent words were presented within each list. The basic experimental word lists from which the random orders were derived are: free recall–Flag, Bird, Chair, Hand, Song, Door, Line; serial recall – Coat, Ball, Meat, Page, Head, Star, Tree. Eighty random orders of the basic experimental word lists were constructed. The practice lists were constant for all subjects so random orders of these lists were not needed.

The author tape recorded each random list for each subject at a rate of one word per 2 sec, without emphasis or inflexion. A Meylan stopwatch was used to time the reading of the words.

Thus, both the practice and experimental lists for both tasks for each subject were pre-recorded. During the experimental phase each group participated in two tasks, counterbalanced in presentation. The subject was brought into the experimental room and seated at a table opposite the experimenter. The tape recorder was placed directly in front of the subject.

In Task I (multitrial free recall), the experimenter instructed the subject to remember as many words as possible, in any order, from the list of words the subject would hear from the recorder. After hearing the list, the experimenter instructed the subject to verbalize all the words the subject could remember. The experimenter recorded verbatim each response the subject made. The subject heard the list in a constant order either until he recalled all the words or five trials had elapsed. Since the free and serial lists had to be recorded five times each for each subject, the task of recording was easier using a constant order for each subject (Waugh (1961) conducted experiments in which identical curves resulted under conditions of a constant order vs. a varied order). The intertrial interval varied depending on the amount of time the subject required to verbalize the words he remembered. Playing of the next list began immediately after the subject had completed making his responses.

Task II (serial recall) began immediately upon completion of Task I and was prefaced by a practice serial recall list. The experimenter instructed the subject to listen and remember the words in the exact order in which the subject heard them. As in Task I, the subject was then instructed to verbalize the words he remembered. The subject continued the task to the criterion of one errorless trial or up to a maximum of five trials. The experimenter made a verbatim written record of the subjects' responses as the subject made each response. The experimenter gave each subject one practice trial, using a practice list of different words. The practice lists served the double function of inducing a 'set' to respond to the lists to be learned and breaking the 'set' which had been established during the learning of the previous list. This was built into the design since in a repeated measures paradigm there is the problem of order or carry over effects from one task to another. Each subject in each group had a random order of the same experimental lists so that effects from any unequal difficulty in any particular position would be randomized in the end results.

The instructions for both tasks were read by the experimenter to the subject so that an initial rapport would be instituted and also for the sake of efficiency in task administration. Unpaced recall was used for both tasks. Both tasks were performed individually in a well lighted, relatively soundproof room, free of overt distractions. During the experiment only the experimenter and subject were present in the room.

RESULTS

Using the raw error data, an Index of Relative Difficulty was computed for each subject at each of seven positions for two tasks. Next, the group Index of Relative Difficulty per position for each task was computed. The performance of each grade level is therefore represented as an Index score at each serial position under each experimental task. The serial position curves, expressed as Indices of Relative Difficulty, are presented in Figs 1 and 2.

The curves were compared by multiple discriminant analysis (Overall & Klett, 1972). The actual computation of the data was performed by an IBM Series 370 Model 165 computer at the University of Missouri Computational Services Center Columbia, Missouri. Results of the multiple discriminant analysis performed on Index scores are reported in Table 1.

The multiple discriminant analysis on the free recall task was not significant ($F = 1{\cdot}261$, d.f. = 21, 202) indicating that the mean vectors give no information about group membership (i.e. there is no difference in the shape of the curves).

The multiple discriminant analysis on the serial recall task was not significant ($F = 1{\cdot}301$, d.f. = 21, 202) indicating that the mean vectors give no information about group membership (i.e. there is no difference in the shape of the curves).

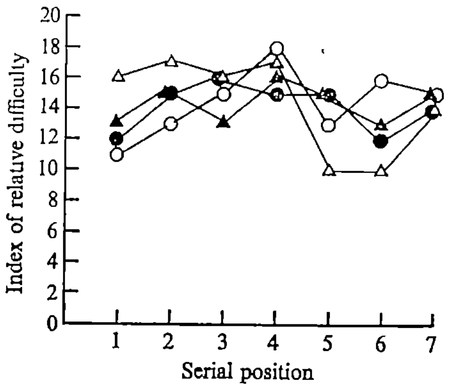


Fig. 1

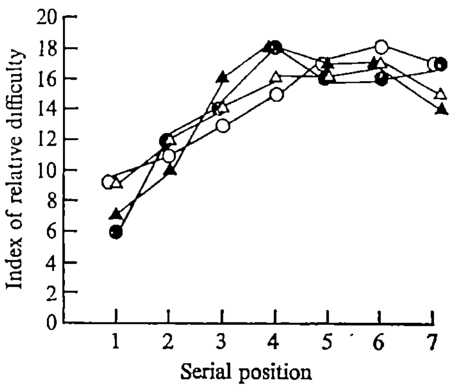


Fig. 2

Fig. 1. Serial position curves for the free recall task by grade level as an Index of Relative Difficulty. △—△, K; ●—●, 1; ▲—▲, 3; ○—○, 5.

Fig. 2. Serial position curves for the serial recall task by grade level as an Index of Relative Difficulty. △—△, K; ●—●, 1; ▲—▲, 3; ○—○, 5.

Table 1. *Discriminant analysis of Index of Relative Difficulty for free recall and serial recall data*

	Group means				*U* statistic		*F* statistic*	
Position	K	1	3	5	Value	d.f.	Value	d.f.
Free recall task								
1	0·16	0·12	0·13	0·11	0·702	7 3 76	1·261	21 202
2	0·17	0·15	0·15	0·13				
3	0·16	0·16	0·13	0·15				
4	0·17	0·15	0·16	0·18				
5	0·10	0·15	0·15	0·13				
6	0 10	0·12	0·13	0·16				
7	0·14	0·14	0·15	0·15				
Serial recall task								
1	0·09	0·06	0·07	0·09	0·694	7 3 76	1·301	21 202
2	0·12	0·12	0·10	0·11				
3	0·14	0·14	0·16	0·13				
4	0·16	0·18	0·18	0·15				
5	0·16	0·17	0·17	0·17				
6	0·17	0·16	0·17	0·18				
7	0·15	0·17	0·14	0·17				

* 1·62 required for 0·05 level.

The Hotelling T^2 analysis (Winer, 1971) was significant ($F = 6{\cdot}859$, d.f. = 7, 73) indicating that the serial recall and free recall curves differ. Related t statistics were computed following the T^2 analysis to determine where the differences occurred. The results are shown in Table 2. Figs 3 and 4 show the serial position curves for the free recall and serial recall tasks expressed as mean errors.

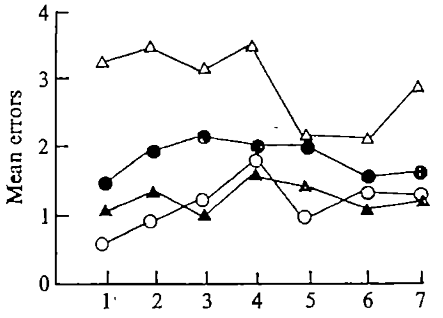


Fig. 3

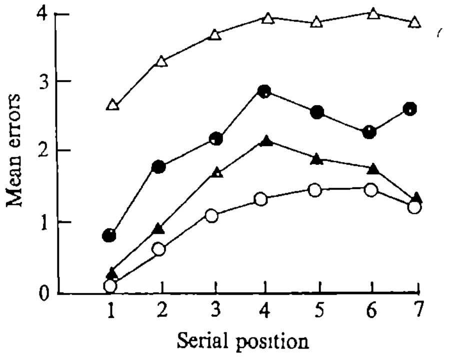


Fig. 4

Fig. 3. Serial position curves for the free recall task by grade level as mean errors. △—△, K; ●—●, 1; ▲—▲, 3; ○—○, 5.

Fig. 4. Serial position curves for the serial recall task by grade level as mean errors. △—△, K; ●—●, 1; ▲—▲, 3; ○—○, 5.

Table 2. *Hotelling* T^2 *analysis of serial vs. free recall curves*

	Means		
Position	Serial recall	Free recall	*t* value
1	0·08	0·13	6·00**
2	0·11	0·15	3·87**
3	0·14	0·15	0·40
4	0·17	0·16	−0·39
5	0·17	0·14	−3·54**
6	0·17	0·13	−4·29**
7	0·16	0·14	−1·42

$T^2 = 43{\cdot}97$; $F = 6{\cdot}8586$**; d.f. = 6, 73; ** Significant at 0·01 level.

DISCUSSION

As indicated in Figs 1 and 2 there were no significant differences in the overall performance of the developmental (grade) levels studied, in terms of mean indices of relative difficulty at each position. Similar findings have been reported by other investigators (Barnett *et al.* 1960; Andrews *et al.* 1954). Noteworthy, however, is that for typical seven-item lists the most difficult item is in position five. The 'flatness' in position five, Fig. 2 is due to the fact that every subject did not perform to a criterion of one errorless trial. As a result, proportionately fewer errors are reflected where a peak in errors is predicted. Although there was a statistically significant difference between the two tasks, there were no significant differences between the grade levels performance on those tasks.

At any rate, in this particular study the interest is primarily in the shape of the serial position curve. Since no significant differences were found in level performance, as shown by a lack of significant differences in the discriminant functions, the shape (bowness and skewness) of the grade level curves are similar.

A significant difference in the tasks essentially corroborates what investigators

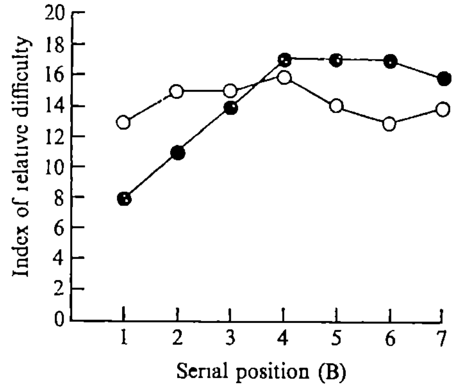


Fig. 5. Serial position curves for the serial recall (●—●) and free recall (○—○) tasks (C) as an Index of Relative Difficulty.

have been finding since the time of Ebbinghaus, i.e. the classic bowed serial position curve as a function of serial position errors or correct responses appropriately plotted. Or, expressed differently, the relative difficulty of learning the word lists has been shown to be a significant function of serial position. Thus far no one has explained, to everyone's satisfaction, what accounts for the serial position curve. There are almost as many hypotheses as there have been studies examining the phenomenon. There is a consensus, however, among most investigators that a multifactor theory in lieu of a single factor theory will likely fully explain the phenomenon (Kling & Riggs, 1972). The person who synthesizes all the various hypotheses accounting for given proportions of the variance for the curve will have solved a problem that has been 'tantalizing' investigators since the experimental analysis of rote verbal learning began. These investigators felt that chronological age (developmental level) could be a significant factor; however, as shown, being in a particular grade level did not make any statistically significant differences as far as performances on the tasks or at positions was concerned.

The significant T^2 is of interest from the standpoint that serial position does act differentially for the free recall and serial recall tasks. An inspection of the plot in Fig. 3 reveals where the differences are to be found. These plots essentially fit the idealized curves for these tasks. That is, in a free recall task there is more of a recency effect than in a serial recall task. Conversely, in a serial recall task there is more of a primacy effect than in a free recall task.

Explicating further, if one were to examine Fig. 4 closely, serial position appears to be more of a factor in performance for grades 1, 3, and 5 than it does for Kindergarten. That is to say, the degree of bowing seems to increase as grade level increases. The writers are suggesting, along with Jensen (1971), that between the ages of 5 and 7 there are significant cognitive or perhaps learning strategy changes occurring. More precisely stated, there is evidence of a change in mode of cognitive functioning during the ages 5–7, which suggests the 'maturation' of some common fundamental mechanisms. Although this particular study does not provide analytical evidence for that, it is suggestive of something else operating or accounting for more variance than serial position, at the grade level of Kindergarten, which in effect is actually a pre-school age of 5.

Figures 3 and 4 are provided merely as comparative data to illustrate differences in mean errors for each level at each position under both tasks. These results would seem to confirm the assumptions made concerning an increase in retention (or conversely fewer errors and thus more efficiency) with age (Edwards & Scannell, 1968; Jensen & Roden, 1963). It will be recalled that the raw error data were converted into Index scores on which the actual discriminant analysis was performed. These figures illustrate the differences of expressing error data in terms of absolute versus relative indices. It might be noted here that the procedure for computing an Index of Relative Difficulty is somewhat laborious. Although the Index does indeed serve a useful, valid function another method of expressing the position effect, which is what ultimately the Index involves, is: To compute mean percentage of total errors at each position for each subject individually prior to averaging over the group. This yields equivalent weighting by subjects rather than by responses.

In summary, this investigation indicated that the shapes of the serial position curves are comparable across developmental levels, as reflected in the findings of no significant differences in the group performances on the experimental tasks administered.

In light of these findings, it would seem that the serial position effect is a relatively invariant function of rote verbal learning. However, the invariance would seem to be more precisely applicable to the bowness and skewness of the curve, rather than to the beginning portion (primacy). Perhaps further studies investigating the serial position effect should make more thorough investigations of the primary and recency effect, rather than the bowness and skewness, since the primary and recency do not seem to be as invariant as the bowness and skewness.

This study is based on the Author's dissertation, submitted in partial fulfilment of the PhD degree at the University of Missouri, Columbia.

REFERENCES

ANDREWS, T. G., SHAPIRO, S. & COFER, C. N. (1954). Transfer and generalization of the inhibitory potential developed in serial rote learning. *Am. J. Psychol.* **67**, 453–463.

BARNETT, C. D., ELLIS, N. R. & PRYER, M. W. (1960). Serial position effects in superior and retarded subjects. *Psychol. Rep.* **7**, 111–113.

BRAUN, H. W. & HEYMANN, S. P. (1958). Meaningfulness of material, distribution of practice, and serial position curves. *J. exp. Psychol.* **56**, 146–150.

EDWARDS, A. J. & SCANNELL, D. P. (1968). *Educational Psychology: The Teaching–Learning Process.* Scranton, Pennsylvania: International Textbook Co.

GOULET, L. R. (1968). Verbal learning in children: implications for developmental research. *Psychol. Bull.* **69**, 359–376.

HOVLAND, L. I. (1938). Experimental studies in rote learning theory. III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.* **23**, 172–190.

HULL, C. L. (1935). The conflicting psychologies of learning—a way out. *Psychol. Rev.* **42**, 491–516.

JENSEN, A. R. (1962*a*). Temporal and spatial effects of serial position. *Am. J. Psychol.* **75**, 390–400.

JENSEN, A. R. (1962*b*). Is the serial position curve invariant? *Br. J. Psychol.* **53** (2), 159–166.

JENSEN, A. R. (1962*c*). An empirical theory of the serial position effect. *J. Psychol.* **53**, 127–142.

Jensen, A. R. (1971). The role of verbal mediation in mental development. *J. genet. Psychol.* **118**, 39–70.

Jensen, A. R. & Roden, A. (1963). Memory span and the skewness of the serial position curve. *Br. J. Psychol.* **54** (4), 337–349.

Kling, J. W. & Riggs, L. A. (1972). *Woodworth & Schlosberg's Experimental Psychology*, vol. II: *Learning, Motivation and Memory*, 3rd ed., pp. 1017. New York: Holt Rinehart & Winston.

Lepley, W. M. (1934). Serial reactions considered as conditioned reactions. *Psychol. Monogr.* **46** (205).

McCrary, J. W., Jr. & Hunter, W. S. (1953). Serial position curves in verbal learning. *Science, N.Y.* **117**, 3032, 131–134.

McGeoch, J. A. & Irion, A. L. (1952). *The Psychology of Human Learning*. New Jersey: McKay.

Murdock, B. B., Jr. (1960). The distinctiveness of stimuli. *Psychol. Rev.* 67, 16–31.

Overall, J. E. & Klett, C. J. (1972). *Applied Multivariate Analysis*. New York: McGraw-Hill.

Spiker, C. C. (1960). Research methods in children's learning. In P. H. Mussen (ed.), *Handbook Methods in Child Development*. New York: Wiley.

Thorndike, E. L. & Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications, Teachers College, Columbia University.

Waugh, N. C. (1961). Free versus serial recall. *J. exp. Psychol.* **62** (5), 496–502.

Winer, B. J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received* 17 *July* 1974; *revised manuscript received* 29 *October* 1975)

Br. J. Psychol. (1976), 67, 4, *pp.* 579–586
Printed in Great Britain

THE DIFFICULTY OF AN IMPLICATION TASK

By DAVID S. BRÉE
Graduate School of Management, Delft, The Netherlands
AND GEORGE COPPENS
Institute for Cognitive Research, University of Amsterdam

An experiment was conducted to distinguish between two cognitive processing models that attempt to account for the difficulty of a reasoning task. The task requires that subjects select the appropriate cards to test the proposition 'if on one half of the card there is a vowel, then there is an even digit on the other half', when only one-half of each card is visible. Few subjects select the correct cards. Johnson-Laird & Wason (1970) put forward a model suggesting that the difficulty is due to a set for verification, i.e. that subjects select cards that could conform to the proposition rather than those which could disconfirm it. An alternative model (Brée, 1973) attributes the difficulty to a failure to hypothesize, i.e. to a lack of consideration of the possible hidden symbols. Only 2 out of 24 subjects in the experiment failed to conform to a prediction of the latter model, which the former model was not able to make.

A simple but counterintuitive finding is always interesting as it may give rise to an explanation that will correct the underlying concept that led to the inadequate intuition. Such a counterintuitive finding, made by Wason (1968), is that an apparently simple reasoning task, based on a conditional proposition, is too difficult for most university students. Two explanations have now been offered for the difficulty of the task. The first put forward by Johnson-Laird & Wason (1970) attributes the difficulty to a set for verification. More recently Brée (1973) has suggested that the difficulty has two independent causes. One reason is that a conditional proposition is often given a symmetrical interpretation, rather than an asymmetrical one. The other reason is that it is necessary in this task to compare a visible symbol with an implication derived from all the possible cases of a hidden symbol, i.e. to hypothesize about what a hidden symbol could be, draw conclusions based on this hypothesis and compare the conclusion with the reality.

In this difficult task, let us call it Wason's validation task, a subject is shown a collection of objects. Each object has two attributes; only one is visible and the other is concealed, e.g. cards with a letter on one side and a digit on the other. The subject is given a conditional proposition about the objects, e.g. 'If the letter is a vowel then the digit is even', and asked to select just those cards which need to be completely revealed in order for him to determine whether the proposition is true or false. Typically in this task subjects select the vowels and frequently the even digits. It is only with 'realistic' material, e.g. with the proposition 'If the letter is sealed, then it has a 50 lire stamp on it', that the 'logically correct' cards are chosen, e.g. sealed letters and letters without a 50 lire stamp (Wason & Johnson-Laird, 1972).

THE INSIGHT MODEL

In their model Johnson-Laird & Wason (1970) suppose that subjects examine the proposition and retrieve the appropriate truth table. They then proceed according to their degree of insight about testing validity. Subjects with 'no insight' select only those objects which could verify the proposition. Subjects with 'partial insight' select those objects which could verify or falsify and subjects with 'complete insight' select only objects that falsify the proposition. Most subjects fall within the first category, having no insight, and suffer from a 'set for verification'. We will call this the 'insight' model.

To determine the prediction of the insight model, it is necessary to specify the form of the truth table that the subjects have stored for a conditional proposition. It is not clear what this truth table should be. The least we can say is that any semantic interpretation of a proposition of the form 'if p then q' must meet the requirement for strict implication, i.e. a $p\tilde{q}$ (p and not q) object must be disconfirming (Lewis & Langford, 1959, p. 100). So in our example with the proposition 'if the letter is a vowel (p), then the digit is even (q)', a card with a vowel (p) and an odd digit ($\tilde{q}$) would be falsifying. Further we suggest that a pq object should be considered as conforming to the proposition, e.g. a card with a vowel and an even digit would be verifying.

In the validation task very few subjects select cards with odd digits ($\tilde{q}$) which they should do if they have partial or complete insight, as these cards can falsify the proposition if there is a vowel (p) on the hidden side. So most subjects have no insight, and they should select those cards which could verify, i.e. the vowels (p) and the even digits (q). This is generally the case. However, several subjects select only the vowels (p). To account for this Johnson-Laird & Wason introduce a postulate: subjects who have no insight are affected by whether or not they believe that the proposition implies its converse. If a subject does not believe that the rule implies its converse, he ignores the consequent term of the proposition, e.g. the even digits (q), and selects only the antecedent term, i.e. the vowels (p). We find this feature of their explanation intrinsically unsatisfying.

We suggest that it is important to separate the way the subject encodes the proposition from the subsequent processing required to select the cards. In terms of the insight model, whether or not a subject believes the proposition implies its converse should be reflected in different truth tables rather than in different processing considerations. This is despite early results obtained by Johnson-Laird & Tagart (1969) that most subjects have the same truth table for conditional proposition. However, as Taplin (1971) suggested, there are different interpretations of conditional sentences (Schetagne, 1970; Noordman, 1972; Smalley, 1972).

THE STRATEGY MODEL

The 'strategy' model put forward by Brée (1973) tackles this problem by assuming that subjects encode the proposition not as a truth table, but as a relationship between p and q. The two most frequent encodings are:

$p \rightarrow q$, the illative interpretation (Chipman, 1971),

Table 1. *Predicted selections* (√) *of objects on the validation task, according to type of evaluation and strategy used*

Proposition ...	'If p then q'	
Encoding ...	$p \leftrightarrow q$	$p \rightarrow q$
Interpretation ...	Converse	Illative implication
Evaluation		
Of objects:		
pq	Conforming	Conforming
$p\bar{q}$	Disconfirming	Disconfirming
$\bar{p}q$	Disconfirming	Irrelevant
$\bar{p}\bar{q}$	Irrelevant	Irrelevant
Validation		
Of objects ...	$p\ \bar{p}\ q\ \bar{q}$	$p\ \bar{p}\ q\ \bar{q}$
Using strategy:		
A	√ . √ .	√ . . .
B	√ √ √ √	√ . √ √
C	√ √ √ √	√ . . √
Using insight:		
None-not converse	√ . . .	√ . . .
None-converse	√ . √ .	√ . √ .
Partial	√ √ √ √	√ . √ √
Complete	√ √ √ √	√ . . √

e.g. the presence of a vowel (p) requires the presence of an even digit (q)

$p \leftrightarrow q$, the converse interpretation,

e.g. the presence of a vowel (p) requires the presence of an even digit (q), and vice versa.

Subjects then use one of three distinguishable strategies to select objects:

A: select an object if the visible symbol requires the presence of a particular hidden symbol;

B: select any object which could have a symbol (hidden or visible) requiring the presence of *any* other particular symbol;

C: select as in A. Then select any object which could have a hidden symbol that would necessitate the presence of a symbol *other than the symbol already visible.*

For example, subjects interpreting the proposition as illative implication would select a card with an odd digit ($\bar{q}$) on it only if they used strategy C. On the back of such a card there *could* be a vowel (p) which would require an *even* digit (q) to be on the front of the card. But there is an odd digit ($\bar{q}$) on the front of the card, so the card must be selected using strategy C. However, subjects using either strategy A or B would not select this card.

The experiment reported in this paper was designed to test the predictions made by the strategy model. These predictions concern the selection a subject may make in the validation task, given a particular encoding of the proposition, as revealed by an evaluation task. These predictions are shown in Table 1. For example, subjects interpreting the proposition as illative implication should never select a card with an odd digit ($\bar{q}$), no matter which strategy they use. But those interpreting it as a converse, should always select such a card.

For the purpose of testing the strategy model, attention will be confined to those subjects who evaluate the proposition as either a converse or as illative implication and who, in the validation task, select one of the common patterns of cards as shown in Table 1. A strong test of the model would be that all subjects conform to the prediction of the model. A weak test is that the distribution of subjects in a 2 by 2 contingency table of interpretation (converse vs. illative implication) and selection (p, q or all instances vs. just p or p, q, $\tilde{q}$ or p, $\tilde{q}$) could not have arisen by chance.

METHOD

Procedure

Each subject was tested individually. He was shown a sample card with a letter on the left and a digit on the right, and then given one minute to memorize the proposition (in Dutch): '*Als op de ene helft van een kaart een klinker staat, dan staat er een even cijfer op de andere helft*', which may be translated as 'If on one half of a card there is a vowel, then there is an even digit on the other half', and has the form 'if p, then q'. He was then given two tasks to perform: the validation and evaluation tasks. For each task eight cards were presented simultaneously in a random array.

The *validation* task used the eight concealed cards. Subjects decided which cards they would need to have completely revealed to test whether the proposition was or was not true for the eight cards separately.

The *evaluation* task used the eight revealed cards. Subjects decided whether each card conformed to, disconfirmed or was irrelevant to the proposition.

Material

The cards used in the tasks each had a letter on the left half and a digit on the right half. The eight 'concealed' cards had one side covered, which left one symbol visible. The visible symbols were two vowels (A, E), two consonants (K, R), two even digits (2, 6) and two odd digits (1, 5); that is two instances of p, $\tilde{p}$, q and $\tilde{q}$ respectively. The eight 'revealed' cards showed two instances each of pq (A2, I4) $p\tilde{q}$ (E1, U3) $\tilde{p}q$ (G8, R6) and $\tilde{p}\tilde{q}$ (K5, M7). Answer sheets were prepared for each task.

Design and subjects

Subjects received the task in one of two orders: Group VE: Validation, Evaluation; Group EV: Evaluation, Validation.

Subjects were first- and second-year psychology students at the University of Amsterdam who volunteered to take part in this experiment to fulfil a course requirement. They selected the group they went into by choosing the time of their experiment. Sufficient subjects were run until there were 12 consistent subjects in each group. A subject was classified as inconsistent if he failed to make the same decision for both cards on any pair of instances, e.g. for both the A and E symbols as instances of p, on *all* the tasks.

RESULTS

Fourteen subjects were run before 12 consistent subjects were obtained for group VE. Thirteen subjects were needed to obtain 12 consistent subjects for group EV. We consider the data from the 12 consistent subjects in each group as shown in Table 2. As expected, there was no significant difference between the two groups of subjects. This was checked by considering the four logical possibilities in the validation task, i.e. the vowels (p), the consonants ($\tilde{p}$), the even digits (q) and the odd digits ($\tilde{q}$), separately. For each possibility there was no significant difference in the number of subjects selecting the card between the two groups (using the Fisher exact probability test with a probability of 0·05). A similar set of

Table 2. *Number of subjects classified by their pattern of selection in the validation task and their interpretation in the evaluation task*

		Selection in the validation task						
	For interpretation as ...	Converse imp.		Illative imp.			—	
	The strategy ...	A	B and C	A	B	C	—	
	Predicts selection of ...	(p, q)	$(p, \bar{p}, q, \bar{q})$	p	$(p, q, \bar{q})$	$(p, \bar{q})$	$(p, \bar{p})$	Total
Group	Interpretation given in evaluation task							
VE	Converse implication	5	.	.	.	.	.	5
	Illative implication	1	.	1	1	2	1	6
	Other	1	.	.	.	.	.	1
	Total	7	.	1	1	2	1	12
EV	Converse implication	6	.	.	.	.	.	6
	Illative implication	1	.	.	1	1	.	3
	Other	3	.	.	.	.	.	3
	Total	10	.	.	1	1	.	12
Both	Converse implication	11			.		.	11
	Illative implication	2			6		1	9
	Other	4			.		.	4
	Total	17			6		1	24

tests was made for the data from the evaluation task, where for each of the four logical possibilities, i.e. vowel and even digit (pq), vowel and odd digit ($p\bar{q}$), consonant and even digit ($\bar{p}q$) and consonant with odd digit ($\bar{p}\bar{q}$), the subjects confined themselves to two of the three possible evaluations. For example a consonant and an even digit would be evaluated as either disconfirming or irrelevant but never as conforming. So the Fisher exact probability test could again be used, and, as before, no significant differences were found between the two groups. So the data for the two groups were combined.

The results on the evaluation task are similar to those found previously (Schetagne, 1970; Noordman, 1972). That is about half the subjects interpret the 'If p, then q' proposition as a converse implication. Fewer subjects than expected interpreted the proposition as an illative implication, i.e. any card without p is irrelevant. The results on the validation task are similar to those found by Wason (1968), except that only one subject selected only the p cards.

Nineteen of the 24 subjects evaluated the proposition as a converse or as an illative implication and selected a pattern of cards predicted by the strategy model. Only two of these 19, those who evaluated the proposition as illative implication but selected both the p (vowel) and q (even digit) cards, do not conform to the predictions of the strategy model. Thus the model is weakly confirmed (Fischer exact probability test, $P < 0{\cdot}005$). However, the presence of these two subjects prevents the strategy model from receiving strong confirmation.

It is possible to save the strong version by claiming that these two subjects failed to fulfil the *ceteris paribus* conditions. For example that they changed their interpretation during the course of the experiment. However, we have observed the same abnormality elsewhere (Brée & Coppens, 1975) and do not believe that we should resort to such devices.

An unexpected interaction can be observed in the data interpreted using the strategy model. Whereas all 11 subjects interpreting the proposition as a converse implication used strategy A, only three of the eight subjects interpreting it as an illative implication used this strategy. In previous studies in which many subjects select only the p cards this interaction would not have been observed if the strategy model is correct. This suggests that such an interaction may be peculiar to this experiment.

DISCUSSION

The results of this experiment suggest that there is some support for the strategy model, although the model does not offer a complete explanation.

Additional support for the strategy model comes from a related study by Kunst (1971). He took protocols from 12 subjects thinking aloud while attempting the validation task. As reported elsewhere (Brée, 1973), those subjects who selected any of the patterns of cards predicted by the strategy model for the subjects interpreting the proposition as illative implication (see Table 1), specifically stated that they assumed that the proposition did not imply its converse. Also most of the subjects selected a card 'to see if it has a . . . on it'. For example, they selected a card with a vowel (p) to see if there was an even digit (q) on the hidden part. This is precisely in accordance with the simplest strategy (A) of the model.

Having seen that the data provide support for the strategy model, the question arises: could the insight model have coped with this data? The protocol evidence just mentioned can also be interpreted as a set for verification. Indeed this is the interpretation given by Goodwin & Wason (1972) to subjects' post-experimental explanations of why they made their choices. While we prefer the explanation that, for instance, a subject selects a vowel (p) because he wants to check *whether or not* it has an even digit (q) on it, it is also possible to consider this as simply seeking to verify that there is indeed a hidden digit (q). The insight model is also able to account for the selections made by all 19 of the 24 subjects that were used to test the strategy model. However, the insight model can only account for the selections of the two subjects who presented a difficulty for the strategy model by making a somewhat bizarre assumption. These two subjects who selected p (the vowel) and q (the even digit) in the validation task, evaluated the $\bar{p}q$ (consonant and even digit) card as irrelevant. The insight model is able to account for these two subjects by assuming that they interpreted the proposition as implying its converse, despite their evaluation of the $\bar{p}q$ card as irrelevant. Thus although they considered the card with an even digit and consonant as irrelevant to the proposition 'if the letter is a vowel, then the digit is even' they must still be classified as interpreting the rule as implying its converse. In our opinion this is unsatisfactory. So it is the same two subjects who present a difficulty for both the insight model and the strategy model.

A simplification of the insight model would be able to account for the data of all but one of the 19 subjects. This simplification requires that subjects with no insight are *not* influenced by whether or not they believe the proposition to imply its converse. So in the validation task they simply select *any* card that would verify the proposition, without restriction (cf. the no insight condition of the original version). Although this might be a satisfactory solution for this set of data, it is unable to account for subjects who select only the p (vowel) card. While there was only one such subject in this experiment, the proportion is usually higher (Johnson-Laird & Wason, 1970). It remains to be seen whether or not this simplified version will account for a larger sample of subjects under the conditions used in this experiment.

If the strategy model is compared with the two versions of the insight model, one can summarize the situation in this way:

(1) If no subjects select just the p (vowel) card in the validation task, then the simplified insight model is adequate.

(2) If, for those subjects who select either just the p (vowel) card or both the p (vowel) and q (even digit) cards, there is no systematic difference in the way they interpret the proposition, then the original insight model will be supported.

(3) If, on the other hand, those subjects who select just the p (vowel) card in the validation task interpret the proposition as an illative implication, while those who select both the p (vowel) and q (even digit) cards interpret the proposition as a converse implication then the strategy model will be supported.

A priori it seems to us that the first possibility is unlikely and that there will be neither a complete lack of any connexion as required by the second possibility nor a consistent difference as required by the third. An adequate account of such a situation could be achieved by considering a combination of the strategy and simplified insight models. For example, some subjects may act in accordance with the insight model while others act along the lines suggested by the strategy model. A more useful step would be to construct a more fundamental model than either the insight or strategy models.

REFERENCES

Brée, D. S. (1973). The interpretation of implication. In A. Elithorn & D. Jones (eds), *Artificial and Human Thinking*. Amsterdam: Elsevier.

Brée, D. S. & Coppens G. (1975). The importance of context in the interpretation of conditional sentences. Working paper W/75/18, Delft: Interfakulteit Bedrijfskunde.

Chipman, L. (1971). Material and illative implication. *Mind* **80**, 179–193.

Goodwin, R. Q. & Wason, P. C. (1972). Degrees of insight. *Br. J. Psychol.* **63**, 205–212.

Johnson-Laird, P. N. & Tagart, J. (1969) How implication is understood. *Am. J. Psychol.* **82**, 367–373.

Johnson-Laird, P. N. & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychol.* **1**, 134–148.

Kunst, H. (1971). Een analyse van inzicht in een redeneer taak. Report FL. 01.07.71.068. Amsterdam: Instituut voor Cognitie-onderzoek, Amsterdam University.

Lewis, C. I. & Langford, C. H. (1959). *Symbolic Logic*, 2nd ed. New York: Dover.

Noordman, L. G. M. (1972). On the interpretation of conditional sentences. *Heymans Bull.* no. 72–118. Groningen: Psychologische Instituten, Groningen University.

Schetagne, J. P. (1970). L'implication logique et deux erreurs d'inférences. MA thesis: Montréal: Université de Montréal.

Smalley, N. S. (1972). Structure and process in deduction from a rule. Mimeograph. Plattsburgh: State University of New York.

Taplin, J. E. (1971). Reasoning with conditional sentences. *J. verb. Learn. verb. Behav.* **10**, 219–225.

Br. J. Psychol. (1976), 67, 4, *pp.* 587–592
Printed in Great Britain

A COORDINATION OF HAND PREFERENCE AND SKILL REPLICATED

BY MARIAN ANNETT

Department of Applied Social Studies, Lanchester Polytechnic, Coventry

Annett (1970*a*) described 283 subjects classified for hand preference and measured for the movement time of each hand on a peg moving task. A new sample of 804 subjects has been examined and classified as before. The mean differences between hands for each preference group are very similar in both samples.

Hand preference can be regarded as anchored to an objective measure of manual skill. Both preference and skill are continuously distributed variables, not discrete ones. The quest for the 'essence' of left handedness should be replaced by studies of the characteristics of the distributions involved and of the criteria used to identify segments of the distributions.

The scientific study of handedness and its relation with other lateral asymmetries has been greatly handicapped by a lack of consistency in the incidences of left handedness reported. In several papers it has been argued that the inconsistencies are due largely to the group of mixed handers, those who prefer the right hand for some skilled unimanual actions and the left hand for other actions normally performed by a right hander with the right (Annett, 1964, 1967, 1970*a*, 1972).

Mixed hand preferences cannot be attributed to cultural pressures on natural left handers to use the right hand (Gillies, MacSweeney & Zangwill, 1960). When the preference patterns of 2321 young adults were studied by an association analysis (which makes successive divisions of the sample on the basis of the item most highly correlated with all the others at each stage) a large number of preference groups were distinguished (Annett, 1970*a*). There was no evidence of major subgroups but rather that hand preferences vary continuously between consistent right and consistent left.

Studies of hand strength (Woo & Pearson, 1927) and speed (Annett, 1972) have found distributions of differences between the hands which are continuous, unimodal and approximately normal. In two samples of children, the mean differences between hands in movement time were shown to be systematically related to preferences, when subjects were classified into four groups, right- and left-handed writers each subdivided into pure and mixed cases (Annett, 1970*b*; Annett & Turner, 1974). In undergraduates and older schoolchildren observations of hand usage for the 12 actions used in the association analysis and measures of manual speed were available. By coordinating these two sets of data, preferences could be classified into eight subgroups, ordered fairly systematically for relative manual speed (Annett, 1970*a*).

Whether preferences are classified into 2, 4, 8 groups or any other number is arbitrary in a continuous distribution. What is important is to establish whether the continuum of preference can be coordinated reliably with the continuum of relative manual skill. In defining eight preference groups an attempt was being made to identify eight regions of the continuum; this means that the whole could be divided into subgroups by using any one or more of seven possible criteria. The question is whether these divisions would be consistent from sample to sample. If the individuals classified in the eight subgroups have similar mean differences in manual speed from one sample to another, we can be confident first, that hand preference and skill are consistently related and second, that our criteria of preference reliably identify regions of the preference/skill continuum. This paper reports a replication of the original coordination of preference and skill in a new large sample.

METHOD AND SUBJECTS

The observations of hand preference and measures of time for peg moving were made by students examining one another during a laboratory class on lateral asymmetry at the Open University (OU) summer schools on the course, Introduction to Psychology (DS 261) in 1974. Students were asked to copy their results for these and other tasks on a data summary sheet; the sheets were collected at each class and school and sent to the author. About two-thirds of the total number of students registered at the schools sent in a usable data summary. There are many possible reasons for omission from the sample including missing the class, loss of summary sheets, sending incomplete information and evidence of physical disability. There were 398 males, mean age 36·7, S.D. 11·4 years, and 406 females, mean age 35·8, S.D. 7·7 years, available for the present analysis.

Observations of hand preference were made for the 12 actions observed in the previous sample and listed in Table 1. Materials were provided for subjects to try out each action, rather than relying on report of preference. Measures of hand speed were made with the peg moving task used in earlier samples and shown to have a test–retest reliability of about 0·69 (Annett, Hudson & Turner, 1974).

The earlier sample (Annett, 1970*a*) consisted of two subgroups. One was psychology undergraduates who examined one another during laboratory classes at the University of Hull. The second was a group of 12-year-old children attending a large comprehensive school in Hull examined by the author and research assistants. These two groups represented complete sets of the available cases; there was no reason for entry into the sample to be biased by handedness.

RESULTS

Table 1 shows how the preference classes were defined and the percentages of subjects in each class in the original and the new samples. There are no significant differences between the samples in the proportion of subjects in each class. This suggests that omission from the OU sample was independent of handedness.

Table 2 shows the mean differences between hands in peg moving time in each class and in both samples. These findings are represented graphically in Fig. 1. For several classes the mean differences are virtually identical. They are especially close in classes 3–7 inclusive.

The only notable difference between samples occurs for class 2. In the earlier sample, those claiming to sweep, shovel or thread needles in the left-handed fashion were more *dextral* in skill than the pure right handers (class 1). As this unexpected finding was true of both subgroups, the undergraduates and schoolchildren, it was considered probably reliable. In the OU sample, class 2 falls between classes 1 and 3 in relative manual speed.

The roughly linear relation between preference and speed differences is broken in both the Hull and the OU samples by class 5 (right-handed writers performing one of the primary actions with the left hand). As a group, these subjects are more dextral in manual speed than class 4 (those who deal playing cards with the left hand). In the Hull sample, the reversal of trend was due mainly to the schoolchildren but

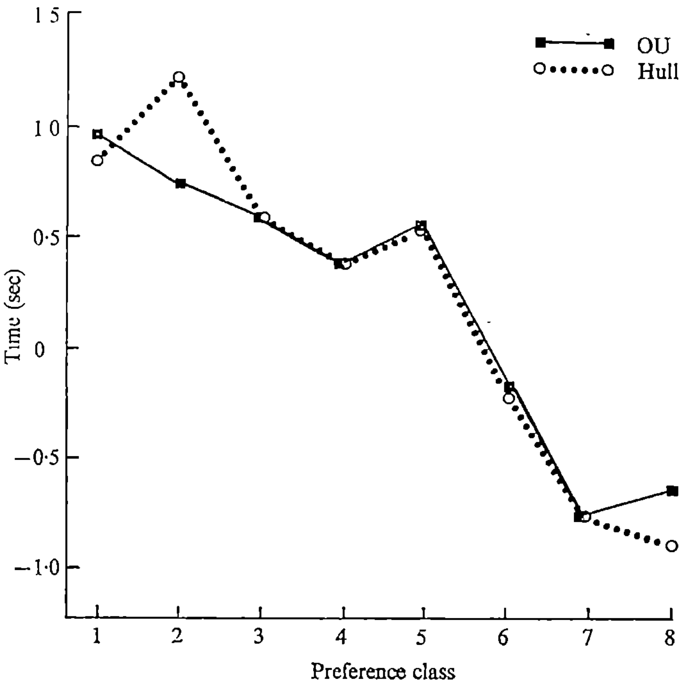


Fig. 1. Mean left hand minus mean right hand peg moving time in hand preference classes.

Table 1. *The actions observed, their grouping into preference classes and the percentages of subjects in the classes*

Actions observed
A Writing*
B Throwing*
C Using a racket*
D Striking a match*
E Cutting with scissors
F Threading a needle (hand doing the guiding whether of the needle or the thread)
G Sweeping with broom
H Shovelling with long handled shovel
I Dealing playing cards
J Hammering a nail into wood*
K Using a toothbrush*
L Unscrewing a jar

Preference classes	Per cent OU sample (n 804)	Per cent Hull sample (n 283)
1 Right for all actions	63·7	58·7
2 Left for any of F, G, H, only	9·4	7·4
3 Left for L and no others except above	7·5	8·8
4 Left for I and no others except above	6·8	8·8
5 Right writing but left for any other primary action	5·0	7·4
6 Left writing but right for any other primary action	2·2	3·9
7 Left for all primary actions but right for any others	1·9	2·5
8 Left for all actions	3·5	2·5

* Primary actions: those most highly intercorrelated (Annett, 1970*a*).

Table 2. *The relation between preference and skill*

Preference class as in Table 1	Left minus right hand mean time on the peg moving task (Means and S.E.M. in sec) Open University (*n* 804)	Hull (*n* 283)
1	0·98 ± 0·03	0·84 ± 0·05
2	0·75 ± 0·09	1·22 ± 0·16
3	0·62 ± 0·09	0·62 ± 0·15
4	0·39 ± 0·08	0·39 ± 0·16
5	0·56 ± 0·12	0·55 ± 0·30
6	−0·16 ± 0·17	−0·21 ± 0·28
7	−0·77 ± 0·22	−0·77 ± 0·36
8	−0·63 ± 0·16	−0·89 ± 0·19

as the numbers from each subgroup in class 5 were small it was not possible to consider this class more closely.

Among OU students there were six subjects, four male and two female, who were first placed in class 5 and then removed to class 8 because they were left handed for *all* actions except writing and gave a history of forced change from sinistral to dextral writing. There had been no comparable cases in the Hull sample.

After removal of these six, there were 40 subjects remaining in the OU class 5 for further analysis. Eleven performed two or more primary actions with the left hand and the mean difference in peg moving time for these subjects was 0·34 ± 0·21 sec. Twenty-nine subjects performed only one primary action with the left hand and the mean difference for this group was 0·64 ± 0·15, close to that of class 3. Only three subjects in class 5 performed more than two primary actions with the left hand; a group too small to suggest that this criterion would be useful. Thus, to identify a group of dextral writers who are more sinistral in skill than other dextrals, we need to take a criterion of two or more primary actions performed with the left hand.

The small difference between the Hull and OU samples in class 8 is not significant statistically. The slightly greater dextrality of the OU class 8 subjects is not attributable to the six dextral writers transferred from class 5. The mean difference in peg moving time for these six subjects was −1·0 sec.

DISCUSSION

The findings for these two large samples are very similar in all essential respects. The practical implications for the classification of subjects in studies of laterality will be considered first before turning to the theoretical implications.

Of the eight preference classes identified in Table 1, the only doubt as to position on the continuum of relative skill arises for class 2. Are subjects who sweep, shovel and thread needles in the left-handed manner truly more sinistral than consistent right handers? It is evident that these subjects are very close to pure right handers and that these actions represent, at most, very trivial departures from pure dextrality. Thus it is doubtful if these actions should be accepted as the only criteria if it were wished to identify subjects with mild but unambiguous sinistral tendencies.

To identify the greatest number of subjects with clear sinistral tendencies and a definite bias to the left hand in skill the safest criterion is left-handed writing. Until recently this criterion was unacceptable because of social pressures against the use of the left hand. To identify the most sinistral of dextral writers, class 5 should be the relevant group but, as defined in Table 1, the group was not the most sinistral in manual skill. Amending the criterion to two instead of one primary action performed with the left hand identifies a group of subjects with the smallest mean difference in peg moving time in favour of the right hand. This newly defined class 5 remains not significantly different from class 4. Since very few subjects were dextral writers and performed more than two primary actions with the left hand we can be confident that individuals who perform *all* primary actions with the left hand apart from writing and give a history of pressure against early sinistral writing should be classified as left handers.

Among left handers those who perform any one of the primary actions with the right hand (class 6) are considerably more dextral in relative skill than those who perform some other actions with the right hand but not primary ones (class 7). It is not certain that class 7 subjects differ significantly from class 8.

In most applications of the present findings to other research samples, it will be necessary to reduce the number of classes. Summarizing the points discussed above we can conclude that the most efficient two group classification in current samples in Britain and other countries without specific pressure against sinistral writing is the single criterion of writing hand. If groups are to be distinguished within dextrals, but not so many as 5, the classes to group together would be 1 and 2, 4 and 5 (as newly defined by two primary actions) and 3 with those now excluded from class 5. Within left handers it would seem more appropriate to group 7 and 8 together than 6 and 7.

Turning now to the theoretical implications of the findings, the most important is that handedness is not a discrete but a continuous variable. Left handedness is often treated as if it were an entity, but one whose exact nature has so far eluded precise definition; the implied aim of research is to identify this mysterious something whose essence psychologists have so far failed to grasp. This amounts to searching for a ghost, though admittedly one that common sense seems to demand. It is a case of psychology being misled by everyday linguistic usage. Left handedness is not a 'thing' in itself, but a region of a continuous distribution of lateral differences. Like many other phenomena treated by threshold models in genetics, and by catastrophe theory in mathematics, it must be thought of as a product of a continuous series of differences which at some point, the threshold or criterion, flips over into the alternate state.

The main contribution of the present replicated findings is that the continuum of preferences can be regarded as anchored to an objectively measured continuum of differences in manual performance. The variability of preferences, which seemed to imply vagueness and unreliability when searching for a single entity, is an intrinsic characteristic. Since preferences can be reliably coordinated with measures of skill, the vagueness evaporates.

The search for a 'true definition of left handedness' has been analogous to the search in psychophysics for a 'true sensory threshold'. The phenomena observed in both cases can be usefully thought of as depending on a continuous and roughly normal distribution. The identification of signals, or left handers, in any region of the distribution is never absolute but subject to misses and false alarms. But the main point of this analogy is that the important research questions concern not the characteristics of left handers as a group, but the nature of the criteria used to identify them and the nature of the distribution as a whole. When an excess of left handers is reported in some subgroup of the population, we need to ask if the excess is due to a generous criterion or to a shift in the distribution. For example, it has been argued that the children of two left-handed parents (Annett, 1974), children with specific language delays (Annett & Turner, 1974) and patients with dysphasias following right unilateral cerebral lesions (Annett, 1975) are subgroups of the population whose laterality distributions lack the shift to the right which is characteristic of samples from the general population in man, but not other species. Many inconsistencies in reports of the incidence of left handedness are attributable to shifts not in the distribution but in the criterion of sinistrality. Some of the uncertainty about the risk of dysphasia following right-sided cerebral lesions in right and left handers has been shown to be due to differences in the criteria of sinistrality applied between studies and between dysphasics and non-dysphasics in the same study (Annett, 1975). Some shifts in the criterion have been due to cultural pressures to dextrality, which make it diffcult to assess whether the low incidences of left handedness in some races and cultures are due to changes in the distribution itself or to changes in the criterion. Studies of relative manual skill would help to solve this puzzle. Recognizing that variations in laterality are continuous rather than discrete entails giving up the search for the 'essence' of left handedness but leads to much more interesting questions about distributions which include very large proportions of the population, the majority of whom manifest as right handers.

I am indebted to the Open University DS 261 course team and to the students who returned information, for the opportunity to collect this data.

REFERENCES

ANNETT, M. (1964). A model of the inheritance of handedness and cerebral dominance. *Nature, Lond.* **204**, 59–60.

ANNETT, M. (1967). The binomial distribution of right, mixed and left handedness. *Q. Jl exp. Psychol.* **19**, 327–333.

ANNETT, M. (1970*a*). A classification of hand preference by association analysis. *Br. J. Psychol.* **61**, 303–321.

ANNETT, M. (1970*b*). The growth of manual preference and speed. *Br. J. Psychol.* **61**, 545–558.

ANNETT, M. (1972). The distribution of manual asymmetry. *Br. J. Psychol.* **63**, 343–358.

ANNETT, M. (1974). Handedness in the children of two left-handed parents. *Br. J. Psychol.* **65**, 129–131.

ANNETT, M. (1975). Hand preference and the laterality of cerebral speech. *Cortex*, **11**, 305–328.

ANNETT, M., HUDSON, P. T. W. & TURNER, A. (1974). The reliability of differences between the hands in motor skill. *Neuropsychologia* **12**, 527–531.

ANNETT, M. & TURNER, A. (1974). Laterality and the growth of intellectual abilities. *Br. J. educ. Psychol.* **44**, 37–44.

GILLIES, S. M., MACSWEENEY, D. A. & ZANGWILL, O. L. (1960). A note on some unusual handedness patterns. *Q. Jl exp. Psychol.* **12**, 113–116.

WOO, T. L. & PEARSON, K. (1927). Dextrality and sinistrality of hand and eye. *Biometrika* **19**, 165–199.

(*Manuscript received* 12 *August* 1975; *revised manuscript received* 15 *January* 1976)

Br. J. Psychol. (1976), 67, 4, pp. 593–607,
Printed in Great Britain

BOOK REVIEWS

Theory and Practice in Behaviour Therapy. By A. J. Yates. New York, London, Sydney, Toronto: Wiley-Interscience. 1975. Pp. xvii + 243. £8.60.

To one who has known Aubrey Yates throughout his career, among his most striking and admirable attributes are the conscientiousness and formality with which he prepares his teaching notes so that they are from the beginning publishable and ultimately published. The present book is the fruit of a series of seminars delivered during 1974 while on sabbatical leave at Queen's University, Kingston, Ontario.

Unlike his previous (1970) work on the same theme, this was not intended to be a compendious handbook. The main intention is to present a viewpoint concerning the nature of and issues within behaviour therapy. That viewpoint is in fact the one already argued forcibly in the earlier volume – that behaviour therapy is neither an application of 'the laws of modern learning theory', nor a set of techniques, but the systematic application of experimental methods in the investigation of behavioural and emotional disorders and their treatment in the single case.

In 1970 great stress was laid on Shapiro's 1950s advocacy of this viewpoint but now Yates traces in fair detail much earlier roots in the nineteenth century writings of Claude Bernard and contemporary parallels in the $n = 1$ advocacy of operant conditioners within the Skinnerian tradition. The inevitable Kuhn is also discussed at some length but with slight profit as, despite reaching the conclusion that psychology is in a pre-paradigm stage, Yates devotes a whole section, with examples, to the importance and use of theories in clinical behaviour therapy.

These arguments are all dealt with in the first chapter. The remaining seven are devoted to an updating of the 1970 position and a review of recent research and practice in relation to a selection of topics. One chapter each is devoted to behaviour analysis (which has advanced greatly in sophistication since 1970), enuresis and stuttering. The next two chapters are concerned with the areas, smoking and obesity, in which behaviour therapy appears to have failed. The fascinating history of systematic desensitization and the collapse of its theoretical basis under experimental attack is then traced and, finally, the new field of biofeedback is reviewed.

Far too many books on behaviour therapy are now being published, but this is one I would have bought if I had not been sent it for review.

H. GWYNNE JONES

Behaviour Modification with the Severely Retarded. Study Group 8 of the Institute for Research into Mental and Multiple Handicap. Edited by C. C. Kiernan and F. Peter Woodford. Amsterdam: Elsevier. 1975. Pp. ix + 332. Dfl. 76.50.

It is not often that a symposium results in a book which is as compulsively readable as this one. Too often we find a collection of ill-assorted papers and bemused commentaries hastily thrown together for the sake of the market. Here, however, C. C. Kiernan & F. Peter Woodford have successfully edited a collection of relevant, intelligent papers into a book which will illuminate this expanding field of research.

The book is organized into three parts: I: ***Fundamental Research***, in which behaviour modification is discussed together with its relation to normal development; II: *Applications*, in which practicalities are exhaustively examined, and III, a final, unusual section in which the contributors look at the therapeutic situation with evaluative and often humorous detachment. This organization renders what could have been a fairly heavy and lengthy meal into three easily digested courses, lacking only a final summing up in which the whole could have been put into perspective. It is a pity that a book as excellent as this should end, not exactly with a whimper, but with a series of disjointed comments, rather than a decisive bang.

The basic philosophy of the book – rethink, re-evaluate, reconsider always in the light of experience – is initiated in the first part by James Hogg, who urges that programmes for the severely subnormal should not be entirely determined by what is considered to be 'normal development', although this may have relevance when dealing with failures. This is refreshing for those who have in the past been asked to justify what seemed to be teaching a child to run before he could walk, or to particularize, to read before he could speak. As Hogg rightly remarks: 'The logical structure of a programme...may be developed independently of the structure of the observed natural development.' This may be recalled later when considering the stimulating paper on communication and language development by W. Yule & M. Berger; the thought occurs: perhaps language acquisition may precede thought? In particular perhaps teaching the pivot-open construction could lead to the child developing the ability to cope with two words in succession, resulting later in the ability to cope with two ideas, i.e. to think constructively and consecutively.

One of the most significant papers in the book is C. C. Kiernan's 'Methodology of behaviour modification', in which the problems of marrying research and practice are explored; many experienced readers will recognize these problems and appreciate the care with which they are discussed. Again and again in this book, one is struck by the relevance of what is written, rooted in the dual role of most of the contributors as theoreticians and practitioners.

Contributors to the section on 'Applications' discuss existing behaviour modification programmes in homes, schools and hospitals. The difficulty of expressing useful therapeutic work in an acceptably 'scientific' framework is well expressed in the paper by Maria Callias & Janet Carr, which illustrates the necessarily nebulous feel about much fieldwork and its successes or failures, together with the complexity of the relationship between therapist and client.

Cliff Cunningham, whose Manchester programme is more ordered and disciplined than many, contributes a confident sensible discussion of the present state of opinion, and attempts to put behaviour modification in its place as only one of possible available techniques. Here, one must agree, up to a point; but I wonder how much of the resistance to behaviour modification among parents and uncertainty among therapists is due to the almost invariable presentation of the technique as some imposed 'scientific' method unrelated to daily life; instead of as an exaggeration of normal parental attitudes and reactions, the only novelty being the intensification and consistency with which these are expressed to children whose attention and comprehension are minimal. A significant and welcome recent change in professional attitudes (medical and educational, as well as psychological) is the recognition that 'parents do feel the need to do something themselves for their child', a thought that, though first expressed nearly 20 years ago, was somewhat forgotten in the era of the psychological personality cult in the early sixties.

Cunningham's question 'whether highly structured teaching approaches with their emphasis on the reduction of uncertainty and a low error rate, teacher control of the learning situation and inflexible objectives are compatible with contemporary conceptions of education' is particularly apposite at this moment; perhaps the answer lies in modifying the contemporary concept; which is, after all, no more sacred than 'normal development'.

Aspects of the practical applications of behaviour modification in schools and hospitals are contributed by Saunders, Jordan and Kiernan, and by Barton; followed by a discussion by Williams and Jackson of an outline training programme for nurses. It is interesting that the only paper which apparently aroused fierce reaction at the time of its delivery was Thorpe's critique of hospital philosophy and its relation to the feasibility of Token Economy systems. As a former hospital worker, this reviewer found much validity in his arguments.

The final section, on the context of programmes, is led by a masterly analysis by Albert Kushlick of the way in which pious hopes and expectations can be transformed into practicalities or as he puts it, how a 'fuzzy' can become a 'performance statement'.

Throughout the book, one is struck by the pertinence of many of the comments in the discussion sections; here the editors have succeeded in capturing the concern and creativity of the participants.

Missing from the book is any discussion of the selection and training of the therapists who, it seems generally agreed, should work in the community with parents, with whom they will interact so crucially. Yule & Berger remark that 'Anyone working with children...ought to have a good grasp of neuropsychology, psychometrics, research design, family and group dynamics, therapeutic dynamics, therapeutic principles and techniques, child development, developmental psycholinguistics, etc.'. We could add to this the skills of the social worker, the teacher, and occasionally the Health Visitor. At the moment, the work is carried out by anyone wishing to research in this area; but if this work is to be justified not only in the particular but in the general, this matter will need as much attention as the other aspects. It seems that an entirely new breed of professional worker will need to be created.

SARAH SANDOW

Psychiatry in Dissent: Controversial Issues in Thought and Practice. By ANTHONY CLARE. Tavistock Publications. 1976. £7.50; paper, £4.00.

The idea that psychiatry is a medical speciality where polymathic pretensions must be allowed to have free reign dies hard. Dr Clare's book is a typical example of this. For in spite of another reviewer's reassurance that it is 'well informed' and 'wide ranging', it is not clear what this book is well informed about or whether its wide and heterogeneous scope is a virtue or, in fact, a defect.

Books can be written for many reasons, both public and private. None of them, however, should be considered good enough to justify repetitious exercises which can only overburden an already saturated market. Dr Clare's book covers a number of topics and with the exception of two chapters (8 and 9) – which could have been better published as two longish articles – the rest either iterate the obvious (chapters 4, 5, 6 and 7) or are frankly weak and conceptually crude (chapters 1, 2, and 3).

Good things first. Chapters 8 and 9 manage to convey well the sense of frustration and uncertainty that many a conventional British psychiatrist feels in his everyday practice. Anybody who is aware of the conceptual, financial (and even political) difficulties bedevilling our psychiatry will certainly agree with Dr Clare's exposition. In fact the Command Paper 'Better Services for the Mentally Ill' (provided that Mrs Castle's introduction is soon forgotten!) conveys a similar, conventional kind of sentiment. So Dr Clare can be reassured on that score that his book represents very well the conventional wisdom and therefore it is not likely to 'interfere with his advancement' as it has been darkly predicted in a recent Sunday newspaper review.

Chapters 4, 5, 6 and 7, on the other hand, dealing as they do with overtrodden aspects of psychiatric practice constitute rather mediocre reviews of topics such as schizophrenia, ECT and psychosurgery. The main confusion in these chapters is that the descriptive and prescriptive contexts are not clearly separated. As a result of this pointless reviews of some aspect of the extant evidence are offered. Pointless on two accounts: firstly, because they unwisely reassure those who wrongly believe that the psychiatric 'truth' is around the corner and its conquest is only a matter of researching away along the same old furrows; secondly, because they fail as a rejoinder to those who have challenged the use of labels such as schizophrenia or of treatment such as psychosurgery and ECT.

'Truth' (whatever that means in this context) could have been better served, in fact, by bringing into the open the real difficulties involved. Firstly, that psychiatry may well share the same regional epistemology of the behavioural sciences which would mean that its natural state is, after all, a 'multiparadigmatic' one (in a Kuhnian sense). Secondly, that an adequate analysis of the present crisis will show that the antipsychiatry movement is not calling into question the amount or the quality of the existing 'evidence' but *what* should count as evidence. In other words, it challenges the present canons of scientificity and their relevance to the interpretation of certain aspects of abnormal human behaviour.

Furthermore, even if at a descriptive level we eventually reach agreement as to what should count as evidence, still that will not provide an answer to their question: Why should these treatments be administered anyway? This because the antipsychiatry group assumes a logical gulf between 'is' and 'ought', and so they are inviting us to play the prescriptive game on a different court and by different rules. The point at issue, therefore, being whether the interventionist ethics of medicine can be extended, without modification, on to psychiatry.

But it is the first three chapters which are the weakest in this already uneven book. An essential requisite to deal with topics such as 'the concept of mental illness' or 'models of mental illness' is to have an adequate knowledge of the background literature in the philosophy of science. The various attempts by Siegler & Osmond, McGuire, Eysenck, Kendell, Kraupl Taylor, Grinker, Leifer and others to sort out this complex area in psychiatry have failed simply because they all have mistakenly believed that the notion of 'model' in psychiatry can be seriously analysed without taking into consideration the conceptual difficulties it has already encountered in the philosophy of science. In fact, it is because these writers confuse the notion of 'model' with notions such as metaphor, hypothesis, structure, theory, analogy, linguistic description, logical class, etc., that their attempts to throw any light upon its use or abuse in psychiatry have so sadly floundered.

But if this type of synchronic, structural analysis of the concept was not possible due to training deficiencies, then at least a diachronic account should have been attempted. An adequate historical study of how the notion of 'medical model' was born in nineteenth century psychiatry would have been welcomed if only because such an account is sorely needed in British psychiatric literature. But Dr Clare is disappointingly timid on both accounts. He does not analyse the problem of the medical model in the context of the philosophy of science nor does he attempt to search for its historical origin. As a result of this he has wasted a very good and expensive opportunity (£4.00 or £7.50) to deal with what is otherwise a fascinating topic.

In view of all this, I would not agree with the claim that this book 'ought to be on the reading lists of psychiatrists in training'. For it will either muddle up their pristine conceptual innocence or (at its best) iterate for them experimental data which can be found, gratis, in any adequate periodical review.

In summary this is a book that falls between two stools: It is too patchy, wide ranging and conceptually confused for the uninitiated (i.e. as a book for John Smith who wants to know what psychiatry is all about). And, at the same time, it is too platitudinous for the initiated who hardly can take yet another London-in-one-day tour through the historical, genetic and clinical common places of schizophrenia, or the statistical data concerning ECT or psychosurgery.

G. E. BERRIOS

Pathways to People. By L. W. DOOB. Newhaven and London: Yale University Press. 1975. Pp. 333. £13.75.

Two indisputably good features of this book are its motive and its topic. Clearly Doob's heart is in the right place and his writing shows a deep concern for the ways in which people can communicate with, and relate to, each other. This is a welcome characteristic in an age when many psychological texts seem to be written by erudite psychopaths for erudite psychopaths. The topic – how people fail or succeed in their attempts to understand each other – is not only itself of central import but Doob strives, at all times, not to trivialize it.

The content of the book is muddled and muddling. The author is concerned both with professional pathways to people (e.g. the study of 'person perception') and with what he terms naturalistic pathways derived from the humanities and from common sense. He rightly feels free to judge the effectiveness of pathways and to suggest strategies. The book has an abundance of pseudo-structure. All paragraphs are numbered and subsets are indexed. Much reference is made to extensive diagrams with reflexive arrows showing how 'ego' judges group 'alters' and also judges their judgements of one another. But peering under all this rather self-conscious organization we find rambling and convoluted discussions of relevant, but arbitrarily chosen, topics – the use of factor analysis as a way of summarizing descriptions of people, the development of skills in person judgement, problems of language in person description, the way in which primary judgements are altered on reflection to secondary judgements, the nature of evidence where judgement of persons is concerned, role theory, the problem of determinism and so forth. Many of Doob's discursive commentaries are not without merit (e.g. he hacks at the apotheosis of the American college student as the prototypical subject, very effectively) but they simply accumulate into a large pile. They do not amount to a theory or even a systematic argument.

Doob is quite bold in his writing tactics but often his boldness can only be given credit of the E for Effort variety. Thus the book is heavily spattered with sentences in italics which represents Doob's notion of what you – the reader – might be thinking about and objecting to at various points in the text. After a while the comments made on the reader's behalf become precious, most readers might find themselves irritated by the attribution to them of reactions which are not theirs and which only serve to remind them how far Doob is not taking their confusions and objections accurately into account.

The book has some virtues then as an 'interesting read'. It is perhaps most significant as a harsh moral lesson for humanist psychologists in search of an alternative psychology. It shows, by its failure, that goodwill and intelligent commentary is not enough.

D. BANNISTER

Safeguarding Psychiatric Privacy: Computer Systems and their Uses. Edited by E. M. LASKA AND R. BANK. New York: John Wiley. 1975. Pp. xv+452. £13.50.

A less elegant and impressive but more accurate title for the work might have been 'The use of computers for record keeping and record use with some comment on what changes, if any, might be required in the laws of the USA to deal with any new problems of confidentiality posed by the use of computers for these purposes'. In other words the reader should be warned that there is little in the book about the uses of computers for purposes other than record keeping, little about technical aspects of keeping computer records secure from unauthorized eyes, and little about the problem of what to put in the records in the first place.

Indeed it is almost possible to see this book as just a non-technical user manual for one particular record system, the Multi-State Information System. Several pages are devoted to completed record forms for this system, and several to descriptions, generated by the computer, of patients whose record forms have been filled in; and most of the book is concerned with this system.

Having said this, it is important to add that it is likely that anyone interested in the use of computers for record keeping would find this book of interest and value. It discusses topics such as the use of problem-oriented case records in psychiatry, evaluation and assessment of community care systems, patient assessment and monitoring, drug monitoring, and, less relevant to NHS practice, ways of keeping up to date with bills for patients. Information scientists and psychiatrists seem to have written most of these sections, and this probably accounts for the lack of concern with the reliability and validity of input data.

About 90 pages in the final section are concerned with legal and ethical issues. Despite the parochiality of some of the subheads in this section, e.g. 'The law in Rhode Island', those considering problems of confidentiality will find a number of interesting ideas and suggestions. One might even go so far as to say that anyone dealing with patients should be given a copy of the 'ten commandments of privacy' on pages 403 and 404.

To sum up, this is not a well-rounded work nor is it a good introduction to the use of computers in psychiatry. It is, however, a book worth reading if you are interested in computer record systems and/or if you want to think about principles and actual legal instruments for dealing with privacy and confidentiality of records.

PHILIP LEY

Drug Treatment of Mental Disorders. Edited by L. L. SIMPSON. New York: Raven Press. 1976. Pp. x+323. $13.50.

Psychopharmacology is one of the boom areas as far as publishing is concerned. The past few years have seen a profusion of texts, single-authored and edited, covering a range of topics both general and specific. It therefore becomes necessary, when reviewing a book such as this, to ask not only whether it has merit when considered on its own terms, but whether or not it does anything which is new or different when considered against the background of literature already available. On both counts, the answer is a clear 'yes'.

L. L. Simpson has assembled a series of 17 chapters, involving 24 different contributors (many of them extremely well-known and productive figures in the field of psychopharmacology), and has managed to present a coverage of modern pharmacotherapeutic techniques which is not only comprehensive and authoritative, but outstandingly readable.

The book is divided into four sections, the first three of which are based on a slightly odd classification of mental illnesses (psychoses, anxiety and affective disorders), the fourth section being devoted to a number of special topics.

In the first section, Seymour Kety presents a succinct account of genetic and biochemical work on schizophrenia and this provides a useful background for subsequent chapters on the phenothiazines (a difficult topic, neatly handled by Jonathan Cole), the thioxanthenes, and the butyrophenones. The third section, devoted to the treatment of affective disorders, also contains chapters reviewing the neuropharmacological background to the drug treatments; there is an excellent coverage of the neuropharmacology of depression, written by Denis Murphy, whilst Baron Shopsin and Samuel Gershon (who have both been active in investigating the use of lithium salts in the treatment of affective disorders) together with Gerard Selzer, cover the neuropharmacology of mania. These two chapters act as background for chapters on tricyclic and monoamine oxidase inhibiting antidepressants, and a very useful chapter (written by the editor) on amphetamines and related stimulant drugs. It is a pity that section two of the book, on the drug treatment of anxiety, did not maintain the format of sections one and three and provide a chapter on the neuropharmacology of anxiety – certainly a difficult chapter to write, but one which would have been both useful and interesting in the context of this volume.

It is heartening to see that the propanediols are given good coverage in a chapter by Frank Berger and also that Ronald Fieve, in considering the uses of lithium salts, also includes rubidium in the title of his chapter. Personally, I expect great things of rubidium salts, but I have had little success in persuading others to take a similar view: Ronald Fieve is a persuasive writer, however, and his chapter will go far to promote the cause.

Amongst the special topics contained in section four two chapters consider the use of drugs in young children and in the aged, whilst Leo Hollister presents a thoughtful coverage of drug side effects and their possible mechanisms.

This book stands out from the mass of literature on psychopharmacology mainly because of the level at which it is pitched. The editor has clearly kept a tight rein on his contributors and the chapters are all of a reasonable length (20–30 text pages, on average) whilst retaining an authoritative flavour. As L. L. Simpson says in his Preface, as the comprehensiveness of a text increases, the likelihood that anyone will read it decreases: this is certainly true when comprehensiveness is confused with exhaustiveness, a confusion which this book successfully avoids.

This excellent, but fairly expensive, text would provide a welcome back-up to lecture courses on psychopharmacology at undergraduate level, but it will also find much wider uses. General practitioners would certainly find that it provides them with a complete updating of their knowledge of available drug treatments for mental disorders, and it should prove invaluable to medical students.

It is clear that we are still only on the threshold of new advances in drug therapy. A book such as this may well be needed every four or five years to remind us where we have reached and to suggest where we might be heading.

F. N. JOHNSON

Sociobiology: the New Synthesis. By EDWARD O. WILSON. Cambridge, Mass. and London: Belknap Press of Harvard University Press. 1975. Pp. ix + 697. £13.60.

E. O. Wilson has been one of the pioneers in the field of sociobiology, and his earlier work *The Insect Societies* is already established as a modern classic study of the behavioural biology of the higher social insects. In this new book he has attempted much more, and overall has succeeded brilliantly. The book provides a comprehensive exposition of the principles of social evolution in animals, and also gives many exciting insights and interpretations into the evolutionary history of the vertebrates and man. Even psychologists familiar with the new concepts of sociobiology cannot fail to be stimulated and challenged by the text; psychologists unfamiliar with the area will have an authoritative and comprehensive source of reference.

The study of animal societies has been undergoing profound changes in the last decade, not only through the field studies familiar to many psychologists, but also through theoretical advances in their interpretation. The main advances have been in the related fields of behavioural ecology, population biology, and sociobiology, all reviewed in the first part of Wilson's book. Through conceptual advances and mathematical simulations it is proving possible to explain much of the behaviour, including social behaviour, of a species in terms of its costs and benefits to the species members. Given a certain subsistence ecology, we can expect – on the basis of time-energy budgets, stability of habitat, and so on – that the species will have certain modal group sizes, feeding strategies, reproductive strategies, and other characteristics. There are costs and benefits involved in social behaviour to conspecifics, too, as well as in adaptations to the physical environment, and here particularly we come to the field of sociobiology. Recent calculations have suggested that Wynne-Edward's theory of group selection as important in the evolution of social behaviour is largely incorrect. Group selection could only be important if a narrow range of parameter values obtain, including a large number of semi-isolated populations with high extinction rates. The evolution of social behaviour must be explained largely if not entirely in terms of individual selection.

Recent insights into possible mechanisms have been achieved by Hamilton, Alexander, Wilson, Trivers and others. A basic problem these theorists have faced, is how cooperative and altruistic behaviours can be selected for under individual selection. By definition, altruistic behaviour implies an individual cost rather than a benefit. Two possibilities however are reciprocal altruism, and kinship selection. In reciprocal altruism, an altruistic behaviour is selected for (via 'altruistic genes' which potentiate an altruistic phenotype) if the altruist can expect other conspecifics to help it in reciprocal circumstances (this will also imply selection against 'cheating'). Kinship selection proceeds from the recognition that an altruist may, by its behaviour, lessen its own chances of survival and hence further reproduction, but thereby enhance the survival of close relatives. As close relatives also share genes in common with the altruist, the altruist's behaviour may increase 'inclusive fitness', in terms of its effect on the overall frequencies of its genes in the species' gene pool.

On the basis of these principles, Wilson proceeds in the second part of the book to overview the main aspects of social behaviour – communication, aggression, territory, dominance, sex, parental care, social symbioses. Much of the descriptive material will be familiar to comparative psychologists, although the explanatory framework may not be.

In the third and final part of the book, Wilson attempts a survey of the social species throughout the animal kingdom. Wilson considers that there are four pinnacles of social evolution – the colonial invertebrates, the social insects, the non-human mammals, and man. The first three show a trend of decreasing sociality, as the coefficient of kinship decreases. In the colonial invertebrates (corals, syphonophores), each organism in the colony comes from the same fertilized egg; the kinship coefficient r is 1, and cooperation and altruism is complete to the extent that the colony simulates an integrated superorganism.

A high degree of social behaviour, involving communal care of young, sterile castes, and division of labour, has evolved at least 11 times in the *Hymenoptera* (bees, ants, wasps), as well as once in the termites. The *Hymenoptera* are haplodiploid species, such that among sisters $r = \frac{3}{4}$ (between mother and daughters $r = \frac{1}{2}$). This high kinship coefficient has enabled high levels of cooperation and altruism to evolve, with only ripples of conflict and competition within the colony. The theoretical advances of sociobiology have been most marked in its explanations of insect sociality, where genetic determination of behaviour is relatively simple, and there are many species available to test theoretical predictions.

Amongst the vertebrates, the highest kinship coefficient (between parent and offspring, or between siblings) is $r = \frac{1}{2}$. As a consequence, the evolution of social cooperation has been greatly tempered by aggressive and competitive tendencies. Amongst the higher mammals, however, enriched communication systems, recognition of individuals, and the integration of learnt experience have been impelling forces towards greater social organization. According to Wilson, this has reached a peak in man, with the conscious recognition of reciprocal altruism and kinship relations, embodied in the exchange rituals and kinship obligations of human societies.

The book includes a valuable and well-illustrated survey of social organization in the mammals. The material here is basically descriptive; compared to the work on insects, the concepts of sociobiology embody a potential rather than a worked-out application to mammalian societies.

Wilson's last chapter is on human evolution and the implications of sociobiology for man. Here it may seem to many psychologists that he adopts too strong a position of genetic influence on behaviours such as conformity to moral values, aggression, or homosexuality, for example. This chapter is based on a much more sophisticated 'biological' approach than is usually found, however, and deserves careful attention. Nevertheless, it does seem to me that Wilson underestimates the complexity and importance of cultural evolution in man, compared to biological evolution. Donald Campbell's recent article (*American Psychologist*, December, 1975) provides an interesting alternative perspective to Wilson's on the importance of cultural evolution in human sociality.

Wilson himself considers that in the next few decades, comparative psychology will be cannibalized by sociobiology and behavioural ecology on the one hand, and neurophysiology on the other. Whether he is correct or not, there is now no doubt of the power and importance of this new field, so elegantly and readably presented in this book.

PETER K. SMITH

Stress Without Distress. By HANS SELYE. London: Hodder & Stoughton. 1974. £2.75.

Dr Selye made his first contributions to the physiology of stress in the 1930s and 1940s with his description of the general adaptation syndrome (GAS), an interrelated series of organic changes involving the hypothalamus, pituitary, adrenocortex and gastro-intestinal tract, which occurred in response to a wide range of environmental stimuli. He has continued to elaborate the detailed physiology of the stress response and particularly its relationship to individual disease processes. Whether or not his theory of a general and non-specific sympatho-adrenomedullary system response to stress is correct, his work has been historically crucial in the development of the whole field of psychosocial stressors and disease.

However, the value of what he has given to the psychological as opposed to the physiological aspect of the question is much more in doubt. Since his interest has been in the mechanism of the GAS he has tended to identify stressors, especially psychological ones, tautologically, as what causes the general adaptation syndrome to be elicited. So even at the WHO sponsored symposium in 1970 on 'Anxiety, Stress and Disease', in which at least three-quarters of the papers were on the nature of psychosocial stressors, Dr Selye's final remarks as Chairman were to call for further work on the physiological measurement techniques.

When he does leave his specialist field to speak on the day-to-day problems of living which preoccupy, excite or bore people, and which constitute the complex phenomenon that those trying to define psychosocial stressors must grapple with, he can offer only a layman's view. But laymen are social and historical creatures. They see the world not from a stance of honest, if ignorant, objectivity, but from the defensive ideologies endemic to their position of oppression or privilege, and since Dr Selye's social round includes the Young Presidents' Club and the Million Dollar Round Table, this layman's perspective is rooted in an ideology which is condescendingly élitist.

If it had not been for the public response to his 1956 publication, *The Stress of Life*, when, as he himself admits with surprise, a few remarks on the 'philosophy of gratitude' attracted the attentions of 'psychologists, sociologists, anthropologists and even clergymen of different faiths', Dr Selye's ideology would have remained private. However, market demand produces supply. With its claim to answer threatening events from Vietnam to Watergate, from the Middle East to outer-space, and with a picture of Dr Selye's handsome head on the dustcover, looking like a well-preserved Albert Schweitzer, this book is obviously aimed at the same group of believers.

Any contemporary work which talks in the first few pages of the 'great laws of nature' should be viewed with suspicion. When the work subsequently embarks on a shuttle service from the molecular to the societal, claiming to identify the basic structures of life at each level, then suspicion should harden into certainty.

The author's first chapter excursion into the physiology of the GAS is merely to fetch some fashionable window dressing intended to set off the contours of his main display, a set of simple-minded moral imperatives guaranteed to transform the interpersonal realities of anyone daft enough to follow them.

The titles of the remaining chapters together with some appropriate quotations are as follows:

Ch. 2. 'Motivation'... 'It is the hunger for achievement which brings joie de vivre' p. 82.

Ch. 3. 'What is the aim of life'... 'To give meaning and direction to life we need a lofty long-range purpose' p. 104.

Ch. 4. 'To earn your neighbour's love'... 'When I was younger I tried really hard but I soon found that try as I might I could not love my neighbour as much as myself' p. 120.

Ch. 5. 'Resumé'... 'Finally do not forget that there is no ready-made success formula which would suit everybody' p. 135.

If one clears away the clouds of pseudo-science the outlines of Selye's philosophy become quite clear even if its capacity to stay erect is doubtful. The author calls it egotistic altruism. In fact it is much more than this involving a social philosophy '...there will always be leaders and followers', a diffuse kind of theology, and an explicit ethical theory. Needless to say, the book is not at all about the psychology of stress.

In principle, though, why not write about ethics? Psychology must have implications for such pivotal concepts as choice, responsibility and guilt. The continuing response of ethical theories to changing scientific knowledge is to be commended. What is irritating about Selye is not that he writes philosophy but that he writes patronizing, ignorant, inconsistent and dishonest philosophy. Patronizing because with every sentence he talks down at the reader. Ignorant because there is little reference to anything which would count as recognized philosophical work. Inconsistent because any beast called altruistic egotism must pull itself apart as soon as it takes its first step in two opposite directions simultaneously. And dishonest because the only chance that such a creature has of survival is for it to remain stationary and its inactivity to be concealed beneath a cloak of appeals to folksy home truths, biological imperatives, world religious views and conservative authoritarian, 'the happy man knows his place' type ideologies.

I would recommend this book only to a limited class of reader. Anyone who wishes not to think but have the illusion of being scientific, wants to have independence of mind, but enjoys being patronized and, fervently wishes that human conflict would cease but refuses to consider the inequalities which generate it, such an individual will enjoy all 136 pages of text. Anyone else will find the stress unbearable.

P. J. O'DONNELL

The Nature of Psychodynamic Interpretation. By N. M. CHESHIRE. Chichester: John Wiley. 1975. Pp. v+229. £5.95.

This excellent book is hard to review and summarize because its arguments are succinct, coherent and without redundancy, although it is by no means terse. Cheshire makes the point that the most appropriate paradigm for the study of human behaviour is not hypothetico-deductive experimentalism but structure modelling as indeed is the case in the natural sciences. With this point of view psychodynamic explanations and interpretations can be regarded as not unscientific. This case is argued with considerable philosophical expertise by an exceedingly careful examination of the meanings of the terms used in the arguments concerning scientific method in psychology and by explication of the assumptions implicit and explicit behind these methods. Although the arguments are directed specifically towards a demonstration of the scientific value of psychodynamic interpretations they are of such general significance that they are relevant to scientific procedures in psychology generally.

Particularly salutary are the attacks made in this book on the ritual obeisance in academic psychology to objectivity. Students are taught (as we are all too well aware at this time of the year) to despise, 'evidence which could be categorized (and therefore condemned) as "anecdotal", "introspective", "subjective", "clinical" or "qualitative": the totem properties, on the other hand, have been "objectivity" (unspecified but played off against the taboo of subjectivity), "quantifiability", "predictability", "repeatability", "publicity" and "experimental control"' (pp. 126–127): jargon terms indeed which prevent thought, create the illusion of scientific rigour, and which are raised regardless of the argument which the evidence is supposed to support. As a culmination of this section there is a neat demonstration of the logical category errors in Sidman's attempt to set out what could constitute scientific data.

There are similar well-executed attacks on operationalism and the perfect circularity of concepts such as operant, stimulus and reinforcement – arguments which it may be possible to meet but which demand at least very careful consideration, and clearly show the logical absurdities in many well-respected expositions, for example that of Staddon, which indeed demands for comprehensibility, Cheshire shows, that the law of non-contradiction be suspended. Nevertheless, the linguistic analysis is not all destructive (only where necessary). It is constructive in that a good deal of the book demonstrates with considerable conviction that it is not illogical to argue back from loose generalizations about behaviour to particular instances, as is of course done in traditional psychoanalytic interpretation. Clever analogues to this position are found in translation from one language to another and in musical interpretation. Indeed, it is suggested that possibly translation is *the* model for all interpersonal understanding.

An outstanding and for this reviewer delightful feature of this book is the elegance of the writing and the literacy of the author. Latin tags abound; there are quotations from Pope, St Augustine, Carroll, Herbert, Virgil and Sheridan, *inter alia*, quotations which genuinely illuminate the meaning not the culture of the author. In summary, this is carefully argued, well-written essay in scientific method, which goes far in allaying the myth of the unscientific nature of psychodynamic theory. No real psychologist should ignore its message.

P. KLINE

Motivation and Dynamic Structure. By R. B. CATTELL AND D. CHILD. London: Holt, Rinehart & Winston. 1975. Pp. xi + 287. £6.50.

The book sets out in full the work on motivation carried out by Cattell and his colleagues over the last quarter of a century; work, incidentally, which is still in progress. The value of this book, therefore, lies in the fact that for the first time we have a coherent account of a large number of investigations which have previously been scattered over journals often difficult of access. In addition, Child, whose writing is known for its clarity, has ensured to as large an extent as is possible that the account is comprehensible to those outside the Illinois group of factorists. Clearly a highly useful book for any psychologists who want to get to grips with the factorial analysis of motivation.

Cattell's studies of the dynamics of personality resulting in the isolation of ergs and sentiments are far less well known than his investigations of temperament. Consequently it is welcome to find in this book that the original basis of this work is carefully discussed. One question that worries any newcomer to this field is (or ought to be) how can one be sure that the putative motivational factors *are* motivational and not temperamental. Put more formally what is the validity of the objective *t* tests and the resulting factors? Cattell & Child are able here to pull all the evidence together. So far it is not as convincing as that for the personality factors, although readers are able to judge for themselves and this is probably one of the most valuable aspects of this book: on almost all aspects of this movitational work, the evidence is presented and discussed in such a way that readers can disagree with the authors and form their own judgements. This contrasts well with other accounts by Cattell. Thus in the 1957 *Personality and Motivation Structure and Measurement* the work was too new for evaluation and in the 1965 *Scientific Analysis of Personality*, lack of space prevented anything more than stating the findings. From this it is clear that this book by Cattell & Child fulfils an important function of allowing psychologists perhaps for the first time to evaluate rationally this unique and fascinating attempt to factor analyse personality dynamics.

However, this function does have one disadvantage. For those psychologists who know this work some of the book is inevitably, as a survey, not new, although this should not put off prospective readers, since the full coherent picture can now be seen. In addition there is some new material and the chapter on structured learning theory where an attempt is made to weave the motivational factors into a theory of learning deserves careful study and much further research.

The book then does what it intended very well. This reviewer has only one objection of any weight. One of the most important features of Cattell's work on motivation (as in all fields) is that it is comprehensive and systematic. It enables theories to be constructed on the basis of the findings. Furthermore, the findings inevitably bear on the vast range of non-factorial, clinical theories of motivation, such as those of Freud, Jung, McDougall and Murray to name but a few. In this book although these names are mentioned, there is little attempt to discuss such relationships. Similarly in the chapter on occupational psychology there is no discussion of other non-factorial work. This omission is a pity, not only because the theoretical import of the work is of great interest but also because it strengthens the entirely fallacious view that factor analysis is a closed world in only psychotic contact with reality.

The book is well produced and strongly bound, although it is marred for an obsessional by a few minor printing errors. It can be strongly recommended as a full and clear account of Cattell's work on motivation, an account, furthermore, which this reviewer has found most useful with students.

P. KLINE

Power. By D. C. McCLELLAND. New York: Irvington. 1975. Pp. xiv + 427. £14.00.

In this book, Dr McClelland presents the results of several years work on the need for power. The general approach follows the same lines as his earlier volume on achievement motivation which appeared in *The Achieving Society*.

The study can be divided into two main sections. First, there is an analysis of the need for power in individuals. The measuring instrument is the Thematic Apperception test. The conclusion is that the need for power can express itself in four ways which normally follow one another in stages during childhood, adolescence and maturity. These four stages are our old

Freudian friends (or enemies), the oral, anal, phallic and genital phases of development, which are extended and applied to the psychology of power. In the first stage the individual is concerned with building up his own power. In stage two he strives for independence from the power of others. The third stage is one of assertion, the desire to wield power over other people. Finally, there is the fourth and mature stage of cooperative group membership. Each of these four stages is associated with the need for power but they are uncorrelated with each other, thus suggesting that the need for power expresses itself exclusively in only one of these four channels.

We are taken next to a consideration of how in different societies power needs are expressed in one or other of these four stages. In India the culture is predominantly oral. In Mexico it is assertive. McClelland is evidently drawn to the Indian philosophy and he describes sympathetically how one of his students adopted the Indian life style.

The final sections of the book are concerned with the measurement of the need for power from samples of the literature in England over the period 1550–1800 and in the United States from 1780–1970. The thesis here is that when the population has high need for power in conjunction with low need for affiliation the nation tends to embark on war or imperialist conquests. It is claimed that the thesis is broadly confirmed, although the theory is weakened because the authors forgot to measure power motivation in the United States immediately before the First World War. This was surely an unfortunate accident and the thesis would have gained in credibility if the omission had been repaired before the results were published.

Apart from this relatively minor omission my chief reservation about the work lies in doubts about the validity of the measuring instruments. Both the TAT on individuals and the scoring of samples of the literature of nations involve some subjectivity and one wonders how easy it would be to replicate the findings. Nevertheless it has to be admitted that McClelland has accumulated an impressive set of results and this will remain an important book on the psychology of power for some time.

RICHARD LYNN

Environmental Interaction. Edited by D. Canter. London: Surrey University Press. 1976. Pp. viii + 374. £7.50.

This is one of a slowly increasing number of books in the area of architectural/environmental psychology. Canter, however, is at pains right away to explain that to label the area of this book as environmental psychology is to imply a development within mainstream academic psychology. This book is concerned with 'the nature of our interactions with our surroundings', which results in material being drawn from areas such as sociology, physics, architecture and planning. Indeed much of the research in this field is pursued more widely by non-psychologists than by psychologists. Therefore much of what Canter and Stringer are saying will be questioned by straight psychologists.

Canter and Stringer, the major contributors, teach in the MSc Environmental Psychology Course at the University of Surrey. From this one may presume that the book is intended as a basic textbook for their course. The other contributors Griffiths, Boyce, Walters and Kenny are respectively a psychologist, physicist, architect and environmental psychology student.

As Canter points out this book can be used in several ways. First, it can be used to give a very good overall view of the development of man–environment studies, and second by selecting only certain chapters, specific areas of research can be studied. It gradually progresses from the micro-environment examining the thermal, acoustic and luminous environments through to the macro-environment examining the spatial environment, the city and the natural environment.

In reading this book the reader will quickly appreciate the complexities of the interactions we have with our environment, the relatively 'early days' of serious research and most of all the lack of serious support for applied research in man–environment studies by psychologists.

One is left with the impression that the environmental psychologists are still trying to establish the credibility of their research in the eyes of mainstream psychologists, and gain respect for it as a 'science'. But as has been pointed out the subject-matter overlaps with many other disciplines such as physiology, geography, sociology, architecture, planning and engineering which one hopes may encourage a move away for its empirically based research which it leans to for its respectability. But despite the 'early days' of research environment psychologists have much to offer designers but unfortunately from my experience I am not sure that designers are really ready for psychology's contribution yet.

Canter and Stringer having covered the wide range of man–environment interactions, Canter finally turns his attention to the role of the environmental psychologist in the design process. Canter having carried out research and lectured in a School of Architecture has a good understanding of the problems facing the designer and is able to outline the design process. He illustrates the constraints facing the designer; the immediacy of decisions; building completion dates; the problems of client and building users being different groups; not knowing the building users' preferences. The criticism of the designer, who is essentially a pragmatist, is that in seeking help from other disciplines they must be able to provide research findings that he can apply to his design.

Since the research psychologist is used to examining theories and formulating hypotheses one of the simplest roles for the psychologist is to examine the design concept and to reiterate the criteria which the building is designed to achieve. Also, because the designer looks to the psychologist to provide him with information on the building user, Canter suggests that much of the research carried out so far is virtually no more than mere opinion poll research. Rather than simply summarize the users' preferences Canter says that psychologists should translate the building users' reactions because in this way he takes his advocacy role seriously and may become a 'professional revolutionary'. It is in this role that the psychologist may argue that if the designer is serious about users' reactions then they should participate in the design process itself. This is something architects have paid lip service to for some time. I hope Canter's suggestions will get through to architects. However, designers and psychologists will first need common ground on which to communicate.

Here, then, is an eminently readable book that will give the reader a very good broad survey of research carried out in man–environment studies in Britain and North America. Unfortunately the book is rather expensive for students, although it does compare favourably with other books of its kind.

J. LOWE

The Sexual Celibate. By DONALD GOERGEN. London: SPCK. 1976. Pp. vi+266. £6.50.

Having defined celibacy as a 'positive choice of the single life for the sake of Christ in response to the call of God' Don Goergen goes on to elaborate the implications of responding to such a call for the sexuality of the Christian man or woman. His work of counselling in this field has no doubt given him many insights, but possibly the inclusion of case material would have helped to convey these. The author prefers, however, a more theoretical approach using the first part of the book to discuss the theology and psychology of sexuality, and the second (shorter) section to discuss practical implications.

Celibacy is seen in the wider context of the call to follow Christ. In his personal response to Christ the individual may pursue his commitment in a state of marriage or celibacy, and Goergen argues that both responses are valid. He draws attention to the fact that, whether married or celibate, the committed Christian is none the less a sexual person, and that in both vocations there is a need to integrate one's sexuality into what it means to be a son or daughter of God.

The joys, problems and pressures that are peculiar to the celibate life are then considered, with topics including friendship and psychological intimacy, solitude and prayer, serving others, and an awareness of one's own personal needs. At times the reader seems a long way away from everyday living and experience, and here case material would have been most valuable. Nevertheless many helpful points are made, and those who delve into this mixture of analytic psychology, Catholic theology, and Goergen's personal insight are likely to find much to interest them.

ANDREW J. GIBSON

The Rat. A Study in Behavior. By S. A. BARNETT. London: The University of Chicago Press. Pp. xiv+318. £12.00.

This is the second edition of a successful work previously entitled simply *A Study of Behavior*. We first ask of any revision 'Is it new?' The answer is certainly 'Yes'. About half the enlarged bibliography refers to work published since the first edition. The format is basically the same except that the physiological component is moved forward in the text, and there is a more extensive coverage of learned aspects of behaviour. The book's strength, as before, is in integrating psychological and ethological work, and attempting to relate our knowledge of the laboratory rat to that of its wild cousins.

Barnett uses the preface to comment on a reviewer of the first edition who showed the book to various colleagues and found 'each had severe doubts about the section on his own subject' while being 'kind enough about the book as a whole'. These comments remain true. It is easy to point to an omission, or debate an interpretation, in any chapter. But the aim of the text is a synthesis and not an encyclopaedic coverage. The flyleaf suggests that the book will be indispensable to researchers. I think not. This is not a handbook of research on the rat. Researchers will read it – to enjoy it – and perhaps learn a little of what their colleagues are up to.

The text is well presented and the prose lucid. Particularly impressive is the author's constant insistence on rigorous definitions, his appreciation of the nature of different explanations, and his precise analysis of what conclusions can legitimately be drawn from a particular experimental technique. The book can be recommended to students for these aspects alone. It is difficult to see its real market, bridging as it does traditional course and subject demarcations. But the first edition was a great success, and all those satisfied customers will find this revision an improvement and a pleasant addition to their library.

D. I. WILLIAMS

Hans Wallach on Perception. By HANS WALLACH. New York: Quadrangle/The New York Times Book Co. 1976. Pp. xviii + 490. $15.00.

This collection of papers by Hans Wallach represents a substantial part of his life's work. He began his long and distinguished career in the field of perception in 1926 at the University of Berlin, as Wolfgang Köhler's assistant. After obtaining his PhD for research on the visual perception of motion and doing some further work on auditory perception (in which he discovered the head-movement principle in the localization of sounds), he emigrated to the USA to join Köhler at Swarthmore College, where he has remained ever since to pursue his research and inspire several generations of students.

One particularly influential outcome of his collaboration with Köhler was their 1944 monograph on figural after-effects. But whereas Köhler contined to search for the electrochemical correlates of perception in terms of the Gestalt brain model, Wallach rejected physiologizing in favour of perceptual psychophysics. As Wallach sees it, the psychologist's main task is to understand the relation between the pattern of stimulation and the resulting perceptual experience. In the case of perceptual constancy, one of several classical problems which has occupied his attention over the years, this has involved him in a search for invariants in the total stimulus field that can be correlated with the invariance of the subject's perception.

His psychophysical approach to perception bears an obvious similarity to that of J. J. Gibson. However, unlike Gibson, Wallach finds it necessary to go beyond the stimulation itself to invoke various supplementary processes (e.g. Gestalt-like 'organizational' processes and 'memory traces') in order to account for perception in situations involving ambiguity and 'cue-conflict'. And unlike both Gibson and the Gestalt psychologists, who inspired Wallach's early work, he makes no attempt to develop a general theory of perception. Instead, he prefers to start with a particular problem (though always with some consideration for its wider theoretical significance) and attempt to solve it by designing a crucial experiment to decide between alternative hypotheses. As he acknowledges in the Preface, it was as Köhler's assistant at Berlin that he learned 'to translate theoretical questions quickly into experiments'. This is clearly what Wallach most enjoys doing in his scientific work, and what he does exceptionally well. Anyone familiar with his work cannot fail to be impressed by the clarity of his experimental reasoning, the ingenuity of his techniques, and the fruitfulness of his hypotheses (e.g. his well-known 'luminance ratio' hypothesis of lightness constancy, and his 'cue-assimilation' hypothesis of perceptual adaptation).

All these virtues are evident in the collection of 34 papers that go to make up this book and the curious student of perception should 'obtain great intellectual satisfaction as he finds many problems of perception analysed and clarified by ingenious experiments and probing conceptualization' (Carl Zuckerman in the Introduction). What he must not expect to find in this book is anything approaching a general statement of the author's views 'On Perception', as distinct from his closely-reasoned experimental analyses of a wide variety of particular problems: the perception of neutral colour, the role of slant in shape perception, the kinetic depth effect, retinal disparity, and so on.... Admittedly the second half of the book, which is devoted to Wallach's extensive series of studies of perceptual adaptation, does tell a more coherent story; but even here the reader must piece the argument together from material contained in 17 separate papers, helped only by a very brief theoretical section in the final paper, 'Theory and more experiments', which turns out to be more about experiments than about theory. Contrary to what is stated in the publisher's blurb, Wallach has not 'distilled' the results of his many years of research: he has simply assembled a substantial number of his papers.

The contents and layout become understandable in view of the author's obvious impatience with attempts to erect general theories of perception, his preference for concepts that keep closely to the experimental facts (so closely that they often serve a descriptive rather than an explanatory function, as the author himself admits, p. 379), and the way in which he has worked, 'pursuing a problem as long as the work yielded worthwhile results and then shelving it until a new idea promised progress'. But it does mean that, excellent though the individual papers are, the book as a whole is scarcely more than a volume of Wallach 'readings'. As such, it will undoubtedly be welcomed by several generations of his own students and by the many other students of perception who admire his work; it should also prove especially useful to postgraduate students about to embark upon research on perceptual learning. But it seems only fair to warn prospective purchasers that of the book's 34 papers, 31 have already appeared in a readily accessible form.

G. W. GRANGER

Sound Perception in Mammals. Edited by R. J. BENCH, A. PYE AND J. D. PYE. London: Academic Press. Pp. 357. £9.80.

This volume is the proceedings of a symposium organized jointly by The Zoological Society of London and the British Society of Audiology held in 1974. Although papers could have been presented on any of the terrestrial mammals, the majority were in fact on man, which makes the title somewhat misleading. A quick glance at the list of contents would indicate that there is probably something to appeal to the psychologist within, although the dust jacket claims to meet the interests of audiologists, otologists, ENT physicians, psychologists, speech therapists, teachers of the deaf, technologists, environmentalists.... However, more detailed reading leads this reviewer to believe that from the psychologist's point of view, this particular whole is rather less than the sum of its individual parts. The contributions included seem to be a collection of largely unrelated papers and no attempt is made to provide any unifying sections. Neither are any of the reactions of the participants themselves to the papers included. If the psychologist interested in audition is seeking for work aimed at illustrating that hearing might be fundamentally important, then he will not find it here, with the exception of the paper by Vourenkoski on the maternal recognition of infant sounds. On the other hand if the psychologist is looking for something to tell him what are the most recent developments in auditory or speech perception, again this is not the book to choose.

Despite the above comments the book does give a good survey of some of the more recent techniques for the assessment and diagnosis of hearing losses. Again, almost inevitably, there is the clinician's tacit assumption that once the threshold change has been specified, the hearing impairment has been defined and there is nothing to indicate that factors such as information-processing difficulties may also be important in deafness. It would seem the main use of this volume to psychologists is to provide a selected summary of the methods available, and results achieved, in the assessment of auditory functioning. The following are some of the contributions which are most likely to be of psychological interest.

On developmental aspects Bosher provides an account of the processes underlying cochlear maturation and the inception of hearing whilst Bench & Mentz describe further developments in techniques for measurement of human foetal auditory response. Brooks gives a useful summary of the importance of middle ear impedance measurements in audiology and Borg & Zakrisson describe some elegant experiments on the activity of the stapedium muscle when subjects are actively speaking and listening to speech.

There are several papers on behavioural techniques for threshold measurement in humans. These include a survey of evoked response audiometry (Solomon), operant audiometry

(St James-Roberts), while Bennett describes a specially designed cradle to measure psychological responses in the neonate. For those more concerned with the animal side, evoked response techniques are discussed by Sales, whilst behavioural audiometry is reviewed and evaluated by Francis. Of more general interest is the discussion by Smith of the possible role of ultrasound in rodent communication, although the importance of this in normal conditions still remains to be demonstrated.

ANDREW ROSTRON

Recall and Recognition. Edited by J. Brown. Chichester: John Wiley. 1976. Pp. x+275. £9.75.

There is much interest in this book both for the student of memory and the philosopher of science. It reveals a newly discovered awareness of the full complexity of that aggregation of processes we call memory. The barrenness of old conceptual tools is illustrated by Postman's contribution, while the uncertainty of the new ideas can be judged from their variety, clearly the study of memory is in a transition stage.

In many ways this book is reminiscent of the Cofer books. It presents a collection of essays which display the state of the art. Cofer's books, though, reported conferences. In Brown's the conference is missing. A suitable theme is provided, but the contributors missed the chance of interacting. They disagree considerably, especially in their interpretation of the distinction between recall and recognition. If you think you know, and can make some generalization regarding relative performance levels (what Warrington refers to as their normal relationship) then study of this book should have a salutary effect.

Perhaps the most constructive idea advanced is that we should set out to develop a natural history of memory. The great variety of tasks set by students of memory suggests that many different processes are being observed. If this is so, it is small wonder common rules cannot be discovered. Cooper & Monk suggest that learning for recognition and learning for recall are separate operations with very limited overlap. The first involves 'list specification and item-to-context association' while the second 'depends mainly on item-to-item associations which are used to build up a retrieval plan'. Note that they consider the subject to be a free agent who develops a plan in order to carry out the required task. A natural history of processes might therefore be less adequate than a compendium of plans and strategies .Recognition of the subject as an agent rather than an object of study is rare, however, in the studies discussed in this book. This is understandable since the apparent readiness of subjects to cooperate by performing stupid tasks such as learning lists of isolated words says little for their independent critical powers.

The futility of interference theory is most clearly shown by its inability to predict the effects of extra-experimental influences. Other lists learned in the laboratory affect the retention of the list being examined, but such experiences as the use of normal language outside the laboratory does not. Clearly those experimenters, such as Bartlett, who kept their sights trained on matters of everyday relevance would never have wasted years exploring phenomena which are artifacts of the laboratory. We should worry that nearly all the contributors to this book would appear to have ignored Bartlett's object lesson. Even his own student.

The erstwhile casual reader should be warned that a quick dip is unlikely to be pleasant. The water is extremely cold. Serious study is required. The various contributions are conceptually far apart and each writer assumes considerable familiarity with his background material. The most impressive common quality is the turgidness of the prose. At one time, I thought the book the culmination of a competition for bad writing. If Bartlett hears of it, he has other reasons to turn in his grave.

H. C. A. DALE

Br. J. Psychol. (1976), 67, 4, *pp.* 609–614
Printed in Great Britain

OTHER PUBLICATIONS RECEIVED

(Inclusion in this list does not necessarily preclude later review)

Adams, J. A. *Learning Controversy: An Introduction.* Homewood, Ill.: The Dorsey Press. 1976. Pp. xi + 383. $13.80.

Adler, G. (ed.) *C. G. Jung. Letters*, vol. 2. Henley-on-Thames: Routledge & Kegan Paul. 1976. Pp. xlvii + 716. £11.50.

Altschule, M. D. *The Development of Traditional Psychopathology. A Sourcebook.* Chichester: Wiley. 1976. Pp. vi + 330. £12.15.

Anastasi, A. *Psychological Testing*, 14th ed. London: Collier-Macmillan. 1976. Pp. xiii + 750. £9.25.

Anderson, R. P. & Halcomb, C. G. (eds). *Learning Disability/Minimal Brain Dysfunction Syndrome.* Springfield, Ill.: Thomas. 1976. Pp. xii + 283. $16.95.

Arnold, W. *Person, Charakter Persönlichkeit*, vol. 5. Munich: Gunter Orzag. 1975. Pp. 464. DM 44; paper, DM 32.

Aronson, E. *The Social Animal*, 2nd ed. Reading: W. H. Freeman. 1976. Pp. viii + 336. £7.90; paper, £2.90.

Baker, A. A. *Comprehensive Psychiatric Care.* London: Blackwell Scientific Publications. 1976. Pp. x + 294. £6.75.

Bauml, M. J. & Bauml, F. H. *A Dictionary of Gestures.* Folkestone: Bailey Bros. & Swinfen (Scarecrow Press). 1976. Pp xxxv + 249. £8.80.

Bernstein, I. N. (ed.). *Validity Issues in Evaluative Research.* London: Sage. 1976. Pp. 134. £2.50.

Berry, P. (ed.). *Language and Communication in the Mentally Handicapped.* London: Edward Arnold. 1976. Pp. viii + 215. £7.50; paper, £3.50.

Bloomfield, H. H., Cain, M. P., Jaffe, D. T. & Kory, R. B. *TM, Discovering Inner Energy and Overcoming Stress.* London: Allen & Unwin. 1976. Pp. xxvi + 292. £4.95; paper, £1.95.

Bolkar, J. A. *Primer for the Nonmedical Therapist.* Chichester: Wiley. 1976. Pp. 131. £4.90.

Bruce, R. D. *Lacandon Dream Symbolism*, vol. 1. *Dream Symbolism and Interpretation.* Perugino, Mexico: Ediciones Euroamericanas, Klaus Thiela. 1975. Pp. viii + 131. $8.50.

Bullough, V. L. *Sexual Variance in Society and History.* Chichester: Wiley. 1976. Pp. xvi + 715. £16.15.

Bunker, B., Pealson, H. B. & Schulz, J. W. *A Student's Guide to Conducting Social Science Research.* New York: Human Sciences Press. 1975. $3.95.

Campbell, S. F. (ed.). *Piaget Sampler.* Chichester: Wiley. 1976. Pp. xiv + 154. £6.10; paper, £2.65.

Chateau, J. *Psychologie des Attitudes Intellectuelles.* Paris: Librairie Philosophique. 1976. Pp. 291. Price not given.

Clarke, A. M. & Clarke, A. D. B. (eds). *Early Experience: Myth and Evidence.* London: Open Books. 1976. Pp. xiv + 314. £6.50; paper, £3.90.

Cleary, A., Mayes, T. & Packham, D. *Educational Technology. Implications for Early and Special Education.* Chichester: Wiley. 1976. Pp. x + 191. £6.95.

Cohen, D. & Wirth, J. (eds). *Testing Linguistic Hypotheses.* Chichester: Wiley. 1976. Pp. xii + 227. £7·65.

Cohen, S. *Social and Personality Development.* London: Collier-Macmillan. 1976. Pp. vii + 227. £3.95.

Colquhoun, W. P., Folkard, S., Knauth, P. & Rutenfranz, Z. J. (eds). *Experimental Studies of Shiftwork.* Opladen: Westdeutscher Verlag. 1975. Pp. 279. Price not given.

Coon, C. S. *The Hunting People.* Harmondsworth: Penguin. 1976. Pp. 459. £1.95.

Cope, R. G. & Hannah, W. *Revolving College Doors. The Causes and Consequences of Dropping Out, Stopping Out and Transferring.* Chichester: Wiley. 1976. Pp. xiii + 190. £8.60.

Davids, A. *Child Personality and Psychopathology: Current Topics*, vol. 2. Chichester: Wiley. 1975. Pp. ix + 256. £10.45.

de Luca, A. J. *Freud and Future Religious Experience.* New York: Philosophical Library. 1976. Pp. 263. $12.50.

Dietrich, R. *Generative Linguistik für Psychologen.* Stuttgart: Verlag W. Kohlhammer. 1976. Pp. 98. DM 22.

Enler, N. S. & Magnusson, D. (eds). *Interaction Psychology and Personality.* Chichester: Wiley. 1976. Pp. xiii + 663. £13.85.

Eysenck, H. J. & Wilson, G. D. (eds). *A Textbook of Human Psychology.* Lancaster: MTP. 1976. Pp. viii + 343. £9.50; paper, £4.95.

Eysenck, H. J. (ed.). *Case Studies in Behaviour Therapy.* London: Routledge & Kegan Paul. 1976. Pp. xi + 355. £7.95.

Falmagne, R. J. (ed.). *Reasoning: Representation and Process.* Chichester: Wiley. 1975. Pp. xii + 275. £10.15.

Ferguson, G. A. *Statistical Analysis in Psychology and Education*, 4th ed. New York: McGraw-Hill. 1975. Pp. x + 529. $12.95.

Formaner, R. & Gurian, A. *Charting Intellectual Development. A Practical Guide to Piagetian Tasks.* Springfield, Ill.: Thomas. 1976. Pp. viii + 119. $10.50; paper, $6.55.

Fuller, B. (in collaboration with E. J. Applewhite). *Synergistics: Explorations in the Geometry of Thinking.* London: Collier-Macmillan. Pp. xxi + 876. £14.00.

Gardiner, R. J. (ed.). *Readings in Human Memory.* London: Methuen. 1976. Pp. 258. £6.00; paper, £2.95.

Geen, R. G. *Personality the Skein of Behaviour.* London: Henry Kimpton. 1976. Pp. xi + 292. £6.90.

Geldard, F. A. *Sensory Saltation. Metastability in the Perceptual World.* Chichester: Wiley. 1976. Pp. 133. £6.15.

Gescheider, G. A. *Psychophysics. Method and Theory.* Chichester: Wiley. 1976. Pp. x + 177. £6.85.

Giora, Z. *Psychopathology: A Cognitive View.* Chichester: Wiley. 1975. Pp. 174. £8.00.

Goldenberg, I. F. *The Helping Professions in the World of Action.* London: Lexington Books. 1975. Pp. xii + 273. £6.75.

Goldman, J. *Becoming a Psychotherapist.* Springfield, Ill.: Thomas. 1976. Pp. 129. $9.50.

Gorsuch, R. L. & Malony, H. N. (eds). *The Nature of Man. A Social Psychological Perspective.* Springfield, Ill.: Thomas. 1976. Pp. ix + 224. $12.75.

Gottschalk, L. A. & Merlis, S. (eds). *Pharmacokinetics of Psychoactive Drugs: Blood Levels and Clinical Response.* Chichester: Wiley. 1976. Pp. xi + 225. £13.70.

Hart, N. *When Marriage Ends: A Study in Status Passage.* London: Tavistock. 1976. Pp. xiv + 277. £6.50; paper, £3.25.

Heath, A. *Ration Choice and Social Exchange: a Critique of Exchange Theory.* London: Cambridge University Press. 1976. Pp. ix + 194. £8.00.

Held, R. & Richards, W. (eds). *Recent Progress in Perception.* Reading: W. H. Freeman. 1976. Pp. 263. £8.50; paper, £3.80.

Hodgkin, R. A. *Born Curious. New Perspectives in Educational Theory.* Chichester: Wiley. 1976. Pp. vi + 131. £5.50.

Hoffman, H. J. *Psychologie und Massenkommunikation.* Berlin: Walter de Gruyter. 1976. Pp. 248. DM 28.

Hollander, E. P. *Social Psychiatry*, 3rd ed. London: Cambridge University Press. Pp. xxiii + 612. £5.25.

Hollander, E. P. & Hunt, R. G. (eds). *Current Perspectives in Social Psychology*, 4th ed. London: Cambridge University Press. 1976. Pp. xiii + 590. £4.75.

Howell, W. C. *Essentials of Industrial and Organizational Psychology.* Homewood, Ill.: The Dorsey Press. 1976. Pp. xii + 275. $5.95.

Jordan, T. E. *The Mentally Retarded.* Columbus, Ohio: Charles E. Merrill. Pp. vi + 622. $18.15.

Jurjevich, R. M. *The House of Freudism.* Philadelphia: Dorrance. 1976. Pp. 228. $10.50.

Kraemer, R. N. (ed.). *The Forbidden Love. The Normal and Abnormal Love of Children.* London: Sheldon Press. 1976. Pp. x + 150. £2.50.

Kreilkamp, T. *The Corrosion of the Sex.* New York: New York University Press. Pp. xvi + 235. $12.50.

Laing, R. D. *The Politics of the Family and other Essays.* Harmondsworth: Penguin. 1976. Pp. x + 119. 50p.

Landy, F. S. & Trumbo, D. A. *Psychology of Work Behaviour.* Homewood, Ill.: The Dorsey Press. 1976. Pp. xiv + 582. $14.50.

Leogrande, E. *Second Chance to Live.* New York: De Capo Press. Pp. (not numbered). $6.95.

Lester, D. *Unusual Sex Behavior. The Standard Deviations.* Springfield, Ill.: Thomas. 1975. $16.75.

Lewis, M. & Rosenblum, L. A. (eds). *Friendship and Peer Relations.* 1975. Pp. xvi + 320. Chichester: Wiley. £11.65.

Liberman, R. P., King, L. W., Derisi, W. J. & McCann, M. *Personal Effectiveness. Guiding People to Assist Themselves and Improve Their Social Skills with Program Guide.* Champaign, Ill.: Research Press. Pp. xix + 169. $5.95, book; $1.95, guide.

Linden, E. *Hopes, Men and Language.* Harmondsworth: Penguin. 1976. Pp. xv + 304. 90p.

Lindgren, H. C. *Educational Psychology in the Classroom*, 5th ed. Chichester: Wiley. 1976 Pp. xviii + 521. £8.00.

Loevinger, J. *Ego Development.* San Francisco: Jossey Bass. 1976. Pp. xx + 504. $17.50.

Loftus, G. R. & Loftus, E. F. *Human Memory. The Processing of Information.* Chichester: Wiley: 1976. Pp. xii + 179. £6·60.

McCubbin, H. I., Dahl, B. B. & Hunter, E. J. (eds). *Families in the Military System.* London: Sage. 1976. Pp. 393. £11.50.

McFie, J. *Assessment of Organic Intellectual Impairment.* London: Academic Press. 1976. Pp. viii + 164. £4.80.

Mack, J. E. *A Prince of Our Disorder. The Life of T. E. Lawrence.* Boston, Mass.: Little, Brown. 1976. Pp. xxviii + 561. $15.00.

McLerachie, W. J., Boyle, C. L. & Moffett, M. M. *Psychology*, 3rd ed. Reading, Mass.: Addison-Wesley. 1976. Pp. xi + 604. Price not given.

Mackenzie, K. D. *A Theory of Group Structures*, vol. I. *Basic Theory;* vol. II *Empirical Tests.* Pp. vol. I, xviii + 265; vol. II, xv + 228. £12.20 set.

McKinney, F., Lorion, R. P. & Zax, M. *Effective Behavior and Human Development.* London: Collier-Macmillan. 1976. Pp. vii + 504. £8.05.

McReady, W. C. & Greeley, A. M. *The Ultimate Values of the American Population.* London: Sage. 1976. Pp. 212. £6.30; paper, £3.50.

Malan, D. H. *A Study of Brief Psychotherapy*. New York: Plenum. 1976. Pp. xiv + 312. $6.95.

Martin, P. W. *Experiment in Depth: A Study of the Work of Jung, Eliot and Toynbee*. London: Routledge & Kegan Paul. 1976. Pp. 275. £7.95.

Mayhew, P., Clarke, R. V. G. & Hough, J. M. *Crime as Opportunity*. Home Office Research Study No. 34. London: HMSO. 1976. Pp. vii + 36. 70p.

Modgil, S. & Modgil, C. *Piagetian Research, Compilation and Commentary*, vol. I. London: NFER Publishing Co. 1976. Pp. 171. £3.60.

Moos, R. H. (ed.). *Human Adaptation*. Farnborough: Lexington Books. D. C. Heath. 1976. Pp. xii + 447. £8.25.

Mungham, G. & Pearson, G. (eds). *Working Class Youth Culture*. Henley-on-Thames: Routledge & Kegan Paul. 1976. £2.75.

Murray, L. B. & Moriarty, A. E. *Vulnerability, Coping and Growth from Infancy to Adolescence*. London: Yale University Press. 1976. Pp. xxiii + 460. £12.00.

Myers, P. I. & Hammill, D. D. *Methods for Learning Disorders*, 2nd ed. Chichester: Wiley. 1976. Pp. xi + 447. £7.75.

Newman, B. & Newman, P. R. (eds). *Development through Life: A Case Study Approach*. Homewood, Ill.: The Dorsey Press. 1976. Pp. xi + 335. $10.95.

Nye, F. I. *Role Structure and Analysis of the Family*. London: Sage. 1976. Pp. ix + 214. £6.85; paper, £4.00.

Overington, I. *Vision and Acquisition*. London: Pantech Press. 1976. Pp. xi + 380. £14.00.

Page, J. D. *Psychotherapy. The Science of Understanding Deviance*. Baltimore, Md.: Williams & Wilkins. 1976. Pp. vii + 510.

Paranjpe, A. C. *In Search of Identity*. Chichester: Wiley. 1976. Pp. vii + 263. £8.30.

Parker, J. G. & Meier, R. D. *Interpersonal Psychology for Law Enforcement and Corrections*. St Paul, Minn.: West Publishing Co. 1975. Pp. xxiii + 290. Price not given.

Patterson, R. L. (ed.). *Maintaining Effective Token Economics*. Springfield, Ill.: Thomas. 1976. Pp. xi + 176. $14.50.

Peto, F. *Cosmic Psychology*. Houston: Yates. 1976. Pp. 130. Price not given.

Pitt, K. J. *Drugs and Drug Rehabilitation*. London: Wildwood House. 1976. Pp. 287. £5.95; paper, £3.25.

Power, P. G. (ed.). *Adolescent Development. Selected Australasian Readings*. Melbourne: Macmillan. 1976. Pp. x + 372. $8.95 (Aus.).

Rama, S., Ballentine, R. & Ajaya, S. *Yoga and Psychotherapy. The Evolution of Consciousness*. Glenview, Ill.: Himalayan Institute. 1976. Pp. xxvii + 332. Price not given.

Raphael, B. *The Thinking Computer. Mind Inside Matter*. San Francisco: W. H. Freeman 1976. Pp. xiii + 322. £9.00; paper, £4.40.

Raven, B. H. & Rubin, J. Z. *Social Psychology: People in Groups*. Chichester: Wiley. 1976. Pp. xx + 591. £7.50.

Reynolds, D. K. *Movita Psychotherapy*. London: University of California Press. 1976. Pp. ix + 242. £7.65.

Reynolds, V. *The Biology of Human Action*. Reading: W. H. Freeman. 1976. Pp. xvi + 269. £6.20; paper, £2.95.

Robinson, D. *From Drinking to Alcoholism: A Sociological Commentary*. Chichester: Wiley. 1976. Pp. vii + 211. £6.75; paper, £2.95.

Rosenzweig, M. R. *Biologie de la Mémoire*. Paris: Presses Universitaires de France. 1976. Pp. 162. Price not given.

Saegert, S. (ed.). *Crowding in Real Environments*. London: Sage. 1976. Pp. 126. £2.50.

Sampson, E. E. *Social Psychology and Contemporary Society*, 2nd ed. Chichester: Wiley. 1976. Pp. 567. £7.50.

Schafer, R. *A New Language for Psychoanalysis.* London: Yale University Press. 1976. Pp. xii + 394. £12.00.

Schatzman, M. *Soul Murder: Persecution in the Family.* Harmondsworth: Penguin. Pp. xiv + 93. 70p.

Scitovsky, T. *The Joyless Economy.* New York: Oxford University Press. 1976. Pp. xvii + 310. £4.95.

Shwinger, P. *Wage Incentive Systems.* Chichester: Wiley. 1976. Pp. xvi + 200. £11.00.

Sokolov, E. N. & Vinogradova, O. S. (eds). *Neuronal Mechanisms of the Orienting Reflex.* Chichester: Wiley. 1976. Pp. viii + 302. £10.80.

Søvik, N. *Developmental Cybernetics of Handwriting and Graphic Behaviour.* Oslo: Universitetsforlaget. 1976. Pp. 373. £9.75.

Spitzer, R. L. & Klein, D. F. (eds). *Evaluation of Psychological Therapies.* London: The Johns Hopkins University Press. 1976. Pp. xiii + 314. £11.90.

Tedeschi, J. T. & Lindskold, S. *Social Psychology.* Chichester: Wiley. 1976. Pp. xiii + 705. £10.95.

Thibault, J. & Walker, R. L. *Procedural Justice. A Psychological Analysis.* Chichester: Wiley. 1976. Pp. viii + 150. £6.10; paper, £3.00.

Thompson, R. F. *Progress in Psychobiology. Readings from Scientific American.* Reading: W. H. Freeman. 1976. Pp. viii + 392. £9.70; paper, £4.50.

Tizard, J., Moss, P. & Perr, J. *All Our Children. Pre-school Services in a Changing Society.* London: Maurice Temple Smith. 1976. Pp. 252. £5.00; paper, £2.50.

Tuckett, D. (ed.). *An Introduction to Medical Sociology.* London: Tavistock. 1976. Pp. 412. £9.00; paper, £4.50.

Tutt, N. *Violence.* London: HMSO. 1976. Pp. 278. £3.15.

Tyrer, P. *The Role of Bodily Feelings in Anxiety.* London: Oxford University Press. 1976. Pp. xv + 128. £6.50.

Vannier, L. *La Typologie et ses Applications Thérapeutiques.* Paris: Doin Editeurs. 1976. Pp. 536. 92.00 Fr.

Villard, K. L. & Whipple, L. J. *Beginnings in Relational Communication.* Chichester: Wiley. 1976. Pp. xi + 275. £3.70.

Vurpillot, E. *The Visual World of the Child.* (Translated by W. F. C. Gilham.) London: Allen & Unwin. 1976. Pp. 372. £9.50.

Walrond-Skinner, S. *Family Therapy. The Treatment of Natural Systems.* Henley-on-Thames: Routledge & Kegan Paul. 1976. Pp. viii + 164. £5.50; paper, £2.74.

Wapner, S., Cohen, S. B. & Kaplan, B. (eds). *Experiencing the Environment.* New York: Plenum Press. 1976. Pp. vii + 244. $15.95.

Weiner, I. B. (ed.). *Clinical Methods in Psychology.* Chichester: Wiley. 1976. Pp. xvi + 678. £15.50.

Welford, A. T. *Skilled Performance: Perceptual and Motor Skills.* Glenview, Ill.: Scott, Foresman. 1976. Pp. xiv + 200. $5.95.

Westerlund, H. B. *Agression, Anxiety and Defence.* Lund: CWK Gleerup. 1976. Pp. vi + 192 + lxxvi. Price not given.

Whittaker, J. O. *Psychology and Modern Life.* New York: Sciences Press. 1976. Pp. 416. $9.95.

White, R. B. & Gilliland, R. M. *Elements of Psychotherapy: the Mechanisms of Defence.* New York: Grune & Stratton. 1975. Pp. xi + 186. £6.25.

Winick, M. *Malnutrition and Brain Development.* London: Oxford University Press. 1976. Pp. xv + 169. £5.00.

Wolfson, J. *Personality and Learning,* vol. 2. Sevenoaks: Hodder & Stoughton with Open University Press. 1976. Pp. viii + 360. £6.50; paper, £3.30.

Wolman, B. B. *The Therapist's Handbook. Treatment Methods of Mental Disorders.* New York: Van Nostrand Reinhold. 1976. Pp. xvi + 539. £14.70.

Wolman, B. B. (ed.). *Between Survival and Suicide.* Chichester: Wiley. 1976. Pp. ix + 195. £9.80.

Wright, R. D. *Understanding Statistics.* London: Harcourt, Brace Jovanovich. 1976. Pp. xii + 500. £6.70.

Zinberg, N. F., Boris, H. N. & Boris, M. *Teaching Social Change* London: The Johns Hopkins University Press. 1976. Pp. xii + 252. £9.05.

EDITORIAL ACKNOWLEDGEMENT

The Editor, Assistant Editor and Co-editors wish to acknowledge the help they have received in the past year from the following persons who have acted as referees for papers submitted for publication in the *British Journal of Psychology*. (June 1975–May 1976)

MRS D. ADLAM
PROF. J. ANNETT
DR M. ANNETT
DR W. S. ANTHONY
MR M. ARGYLE
PROF. J. W. ATKINSON
DR A. D. BADDELEY
DR D. BANNISTER
MR P. BARBER
DR D. BARTRAM
PROF. A. BAUMEISTER
DR N. BEASLEY
DR J. G. BEAUMONT
DR J. BENJAFIELD
PROF. D. E. BERLYNE
PROF. B. BERNSTEIN
DR P. BERRY
DR D. E. BLACKMAN
DR J. E. BLUNDELL
DR R. A. BOAKES
PROF. W. F. BODMER
DR T. G. R. BOWER
PROF. J. M. BRENER
DR D. E. BROADBENT
PROF. D. B. BROMLEY
PROF. J. BROWN
DR R. BROWN
DR W. P. BROWN
PROF. J. S. BRUNER
DR P. BRYANT
DR G. CLARIDGE
DR C. J. COLBOURN
DR W. P. COLQUHOUN
PROF. M. COLTHEART
PROF. K. CONNOLLY
DR R. CONRAD
DR M. COX
PROF. F. I. M. CRAIK
DR R. CROMER
DR H. C. A. DALE
DR D. R. DAVIES
DR G. M. DAVIES
DR A. DICKINSON
DR S. DIMOND
DR M. DONALDSON
DR G. EDMUNDS
DR J. A. C. EMPSON
PROF. D. R. ENTWISLE
PROF. H. J. EYSENCK
DR M. W. EYSENCK
DR B. J. FELLOWS
DR J. M. FINDLAY
DR G. H. FISHER
DR K. A. FLOWERS
MISS H. FRANCIS
DR F. FRANSELLA
DR C. O. FRASER
DR N. H. FREEMAN
PROF. W. D. FURNEAUX
PROF. H. FURTH
DR A. GALE
DR I. E. GORDON
DR J. A. GRAY
DR D. GREEN
PROF. J. GREENE
MISS G. GREER
DR V. H. GREGG
PROF. R. L. GREGORY
PROF. J. P. GUILFORD
PROF. R. HABER
DR M. S. HALLIDAY
DR V. HAMILTON
MR E. J. HAMMOND
DR R. HARPER
DR L. R. HARTLEY
DR P. HARZEM
DR S. H. A. HENLEY
DR B. M. HERMELIN
PROF. P. HERRIOT
PROF. R. T. HERRNSTEIN
PROF. E. R. HILGARD
PROF. C. HINDLEY
DR I. HINDMARCH
PROF. J. HOCHBERG
DR G. R. J. HOCKEY
DR J. HOGG
PROF. D. HOLDING
DR R. I. HORRELL
PROF. W. H. N. HOTOPF
DR W. J. HOUSE
PROF. L. HUDSON
PROF. I. M. L. HUNTER
DR C. HUTT
PROF. S. J. HUTT
MRS S. HUXTABLE
PROF. G. JAHODA
PROF. A. R. JENSEN
DR P. N. JOHNSON-LAIRD
DR A. R. JONCKHEERE
DR S. JONES
DR R. B. JOYNSON
DR Z. KAMINSKA
DR D. C. KENDRICK
DR R. A. KENNEDY
DR C. G. KIERNAN
DR M. KINSBOURNE
DR P. KLINE
DR C. R. LEACH
DR E. C. LECHELT
PROF. P. M. LEVY
PROF. B. LEWIS
DR C. LOMAX
DR G. LOWE
PROF. A. R. LURIA
PROF. R. LYNN
MR M. McASKIE
MR J. G. McCOMISKY
PROF. P. McEWEN
DR A. McGLAUGHLIN
DR B. O. McGONIGLE
DR H McGURK
DR J. MILLER MAIR
DR W. A. MATTHEWS
PROF. A. E. MAXWELL
DR R. MEDDIS
DR S. MILLAR
DR D. C. MITCHELL
PROF. P. J. MITTLER
DR I. MORIN
DR P. E. MORRIS
DR J. MORTON
DR B. B. MURDOCK, JR.
DR R. S. NICKERSON
DR W. ORWIN
PROF. A. PAIVIO
DR G. PASK
DR J. PATRICK
DR K. E. PATTERSON
DR J. PICKERING
DR P. M. A. RABBITT
DR S. RACHMAN
DR A. RADLEY
DR J. RADFORD
DR J. T. REASON
DR R. A. RESCORLA
DR G. RICHARDSON
DR J. O. ROBINSON
DR H. ROSS
DR A. B. ROSTRON
MR J. ROWAN
MR A. N. ROYSE
MR D. A. ROUTH
DR P. A. RUSSELL
DR J. RUST
MRS S. SANDOW
DR R. SERPELL
DR P. SEYMOUR
DR T. SHALLICE
DR L. H. SHAFFER
DR M. H. SHELDON
MR M. R. SHERIDAN
DR D. SIDDLE
MRS M. SHIELDS
DR M. SIME
DR C. H. S. SINGLETON
PROF. W. SLUCKIN
DR D. P. SPENCE
DR G. M. STEPHENSON
DR G. THOMSON
DR R. THOMSON
DR M. TRIESMAN
PROF. P. H. VENABLES
PROF. M. D. VERNON
DR E. WARRINGTON
DR P. C. WASON
DR J. M. WILDING
MR R. L. WILKES
PROF. O. L. ZANGWILL

Oxford Books for Students

Principles and Methods of Social Psychology

Edwin P. Hollander

This third edition has been entirely rewritten and updated. One of the new features is the use of principles of social psychology as headings throughout the book which serve as helpful study guides for the student. Sixty-two cases are included to illuminate particular points and to encourage discussion. Third edition illustrated £5.25

Current Perspectives in Social Psychology

Edited by Edwin P. Hollander and Raymond G. Hunt

For this new edition twenty-three new selections have been added. Most concern current social issues, including social change, aggression, vandalism, sex and institutional roles, racism and prejudice, non-verbal communication, civil liberties, participatory leadership, and international conflict. Fourth edition paper covers £4.75

Emotional Problems of Adolescents

J. Roswell Gallagher and Herbert I. Harris

This is a guide for parents, teachers, priests and others in understanding normal adolescents and helping them in their problems. It has now been completely revised, taking account of social changes which have taken place in the decade since the second edition was published. The chapters on 'The impact of adults upon adolescents', 'Adolescents' sexuality', 'The severe disorders', and 'Alcohol and other drugs' have been entirely rewritten. Third edition £4.95

Child Development and Education

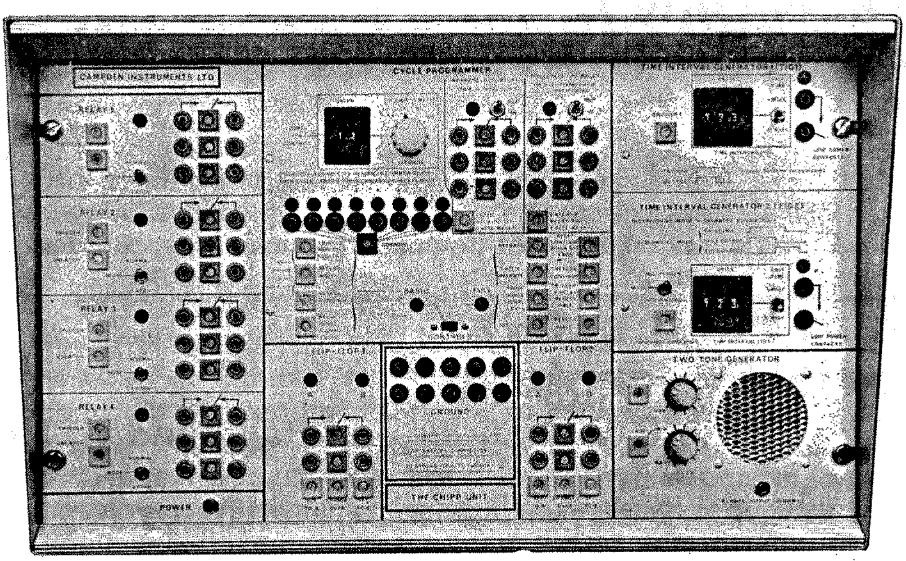
A Piagetian Perspective

David Elkind

This book presents a systematic approach to education from the point of view of child development. The author first provides background information in Piaget's conceptual forerunners and on Piaget's own life and work, and presents some of his most important insights about children and the stages of their cognitive development. He then deals with classroom applications, including methods for determining children's levels of cognitive development and of analysing curriculum materials. £8.25 paper covers £3.25 *forthcoming*

Inspection copies of these books are available to lecturers wishing to recommend them for adoption as required texts. Please write to the University Promotion Section, Oxford University Press (P), Ely House, 37 Dover Street, London W1X 4AH

NEW from CAMPDEN INSTRUMENTS – a powerful and unique package of control logic – THE CHIPP UNIT (Controller for Human Information Processing Psychology)



The CHIPP UNIT has been designed by CAMPDEN in consultation with Dr. John Frisby and Dr. Adrian Simpson of the Department of Psychology in the University of Sheffield who have also written the comprehensive Teaching Manual. The design of the CHIPP UNIT caters ideally for the present day trends in Psychological Experimentation thus superseding existing "Student Laboratory Units"

SPECIAL FEATURES INCLUDE:–

* LOGIC FACILITIES IN A SINGLE INSTRUMENT

* COMPLETE ELECTRICAL ISOLATION BETWEEN THE CHIPP UNIT AND OTHER EQUIPMENT

* L.E.D. INDICATOR LAMPS TO SHOW STATES AT ANY GIVEN INSTANT

* SIMPLE LOGICAL OPERATION BY GROUND SIGNALS

* CLEAR PANEL DESIGN FOR EASE OF OPERATION

* IDEALLY SUITED FOR STUDENT USE AS WELL AS RESEARCH – INCLUDING ANIMAL WORK

The CHIPP UNIT comes with a comprehensive USER'S MANUAL which includes a full description of how it is operated. This Manual includes illustrations on the programming of a number of psychological experiments and is written with the student user particularly in mind so that students should be able to teach themselves how to use it.

For more information on the CHIPP UNIT also on our FISH experimental units write to:–

CAMPDEN INSTRUMENTS LTD.,
186, CAMPDEN HILL ROAD,
LONDON, W8 7TH.

Telephones:– (01) 727 3437
(01) 229 3636

Language and Perception

GEORGE MILLER and P.N. JOHNSON-LAIRD

Language and Perception lays foundations for a new branch of the psychological sciences – psycholexicology, the psychological study of the meaning of words. Although the basic argument is psychological, George Miller and Philip Johnson-Laird also draw on current work in artificial intelligence, linguistics, philosophy, and social anthropology. Their closely argued and lucid treatise will stimulate specialists in many fields to question their assumptions and broaden their thinking about semantic problems. **£12.50 net**

Published in the USA by Harvard University Press

CAMBRIDGE UNIVERSITY PRESS

The Psychology of Sex Differences

Eleanor Emmons Maccoby and Carol Nagy Jacklin

In what ways and to what extent do boys and girls, men and women differ in social behaviour, intellectual abilities, and motivations? This book answers the question in a systematic analysis and interpretation of all reported research findings in such areas as perception, cognition, self-concept, activity level, sociability, aggression, competition, dominance, modelling and socialization, and genetic factors. £7

Human Stereopsis

A Psychophysical Approach

W. Lawrence Gulick and Robert B. Lawson

The authors here describe and relate a decade of research centred on a psychophysical analysis of depth perception. No formal theory is put forward, but the authors have not hesitated to advance a number of theorems, each of which has grown out of and is sustained by the results of their own experiments. £11.75

Oxford University Press

REVUE DE PSYCHOLOGIE APPLIQUEE

Publication Trimestrielle

Directeurs: Dr P. Pichot et P. Rennes

2ème Trimestre 1976 – Volume 26 – No. 2

SOMMAIRE

H. Santucci et M. Verba-Rad – Représentation graphique du bonhomme vu de dos chez l'enfant de 5 à 7 ans.

M. Morval et J. L. Laroche – Constance du dessin de famille.

B. Ducarne de Ribaucourt – Sémiologie de l'aphasie.

Johannes C. Brengelmann – Portrait, notice biographique et bibliographique.

Rédaction et Administration:
48, avenue Victor Hugo – 75783 PARIS CEDEX 16
Abonnements (un an) France: 45F – Étranger: 55F
Le numéro France: 12F – Étranger: 12F
Règlements au nom des
Éditions du Centre de Psychologie Appliquée
Numéro spécimen sur demande

British Journal of Educational Psychology

Vol. 46, Part 3

November 1976

The British Journal of Educational Psychology is concerned mainly with reporting the results of experimentally based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry.

The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words).

There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists.

Subscriptions (£6 a year, or £2 per issue) should be sent to: Scottish Academic Press Ltd, 33 Montgomery Street, Edinburgh EH7 5JX.

CONTENTS

A. E. Bell, M. A. Zipursky and F. Switzer: Informal or Open-Area Education in Relation to Achievement and Personality

A. A. Coulson: The Attitudes of Primary School Heads and Deputy Heads to Deputy Headship

R. F Kempa and G. H. Hodgson: Levels of Concept Acquisition and Concept Maturation in Students of Chemistry

J. Nisbet and J Welsh: Prediction of Failure at University – or Failure of Prediction?

R. J. Lang and K. A. Ryba: The Identification of Some Creative Thinking Parameters Common to the Artistic and Musical Personality

E. A. Lunzer, J. E. Wilkinson and T. Dolan: The Distinctiveness of Operativity as a Measure of Cognitive Functioning in Five-Year-Old Children

E. A. Lunzer, T. Dolan and J. E Wilkinson: The Effectiveness of Measures of Operativity, Language and Short-Term Memory in the Prediction of Reading and Mathematical Understanding

M. A. R. Townsend and B. Keeling: An Investigation of Level I and Level II Cognitive Processes in the Learning and Recall of Factual and Inferential Information

T. Lee, P. Harzem and T. R. Miles: The Effects of Pictures on Learning to Read

G. O'Neill and G. Stanley: Visual Processing of Straight Lines in Dyslexic and Normal Children

Research Notes

A. W. Young and J. McPherson: Ways of Making Number Judgements and Children's Understanding of Quantity Relations

J. Ward and D. Knowles: A Note on the Use of Peel's Logical Game for Studying Conditional Reasoning in Children

J. Adams-Webber and E. Mirc: Assessing the Development of Student Teachers' Role Conceptions

Book Reviews

Publications Received

SCOTTISH ACADEMIC PRESS LIMITED
33 MONTGOMERY STREET, EDINBURGH EH7 5JX

The British Journal of Criminology

Published quarterly on behalf of the Institute for the Study and Treatment of Delinquency by SWEET & MAXWELL Ltd., North Way, Andover, Hants. SP10 5BE

The purpose and policy of the Journal is to foster and publish researches and studies into the causes and prevention of crime, and the treatment of criminals. Each number of the Journal contains original papers. In addition the Journal publishes general surveys of current developments in various countries. On occasion an individual part of the journal will consist of a symposium dealing with a special subject, with papers from groups of experts on issues not only purely scientific but of public moment at any given time.

There are in-depth notes and comments on practical developments, notably in the field of penal reform, on government or privately sponsored reports, and on new experiments in dealing with delinquent or deviant persons. Finally the books and periodicals section provides a comprehensive review of all significant literature in the field.

Annual Subscription £9.50
Single copies £2.75

Mail Orders: Sweet & Maxwell, Spon (Booksellers) Ltd., North Way, Andover, Hants SP10 5BE

Bookshop: Hammick, Sweet & Maxwell, 116 Chancery Lane, London WC2A 1PP

ARCHIVIO DI PSICOLOGIA NEUROLOGIA E PSICHIATRIA
Anno XXXVII Gennaio–Giugno 1976 Fasc. I–II

SOMMARIO

La formazione nella prospettiva psicosociologica

Editoriale
Rouchy, J. C. Psychosociologie et formation
Laporta, R. Prospettive sociali e politiche della formazione
Manoukian, A. Formazione e organizzazione: un'interpretazione storica. L'istituzione della formazione
Kaes, R. 'On (de)forme un enfant': fantasme originaire, processus et travail de la formation
Loué, L., Les différents courants de la formation et leur analyse critique
Enriquez, E. Formation et intervention psychosociologique
Lévy, A. Le groupe comme instrument de formation: analyse critique des différentes perspectives d'utilisation du groupe dans un processus de formation
Carli, R. Trasformazione e cambiamento
Casnati, C. – Capranico, S. Il lavoro formativo attraverso seminari di dinamica di gruppo in più sessioni: relazione di una esperienza

NOTE E DISCUSSIONI

Bertolino, M., *Recenti contributi sull'efficacia deterrente della sanzione penale*

NOTIZIARIO

Incontri e ricerche (Sacchi, A. – Strigelli, M.) – Congressi e convegni – Varie

RECENSIONI

Benkert, O. – Hippius, H. *Psychiatrische Pharmakotherapie. Ein Grundriss für Aerzte und Studenten* – Helmchen, M. – Hippius, H. *Entwicklungstendenzen biologischer Psychiatrie* – Kisler, K. P. – Meyer, I. E. – Müller, M. – Strömgren, E., *Soziale und angewandte Psychiatrie*, III (C. Ferrio)
Fornari, F. *Genitalità e cultura* (M. V. Saraó Levati)
Goldstein, J. H. – Ghee, P. E. *La psicologia dello humour* (G. Siri)

Redaction – Administration: 1 Largo A. Gemelli – 20123 MILAN (Italy)
Annual subscription: Italy L. 11.000; foreign countries L. 16.000 (or equivalent foreign money)

INDEX OF AUTHORS

PAGE
Adams-Webber, J. 11
Agruso, V. M. Jr 569
Annett, M. 587
Benjafield, J. 11
Berlyne, D. E. 175
Bolton, L. 339
Booth, D. A. 137
Boucher, J. 73
Brée, D. S. 579
Campbell, R. 89
Carroll, D. 367
Chambers, S. M. 31
Clarke, A. M. 243
Coppens, G. 579
Craske, B. 461, 475
Crawshaw, M. 475
Davies, G. 63
Davies, P. 181
Day, R. G. 537
Deregowski, J. B. 23
Dickinson, R. G. 537
Ditkofsky, J. 175
Dominowski, R. L. 231
Donaldson, M. 89
Durndell, A. J. 501
Evans, J. St B. T. 479, 517
Eysenck, M. W. 53
Fraser, C. O. 399
Fellows, B. J. 519
Frigon, J.-Y. 467
Gupta, B. S. 47
Hagan, P. J. 359
Hamilton, V. 191
Harris, M. 419
Hinton, J. 461
Henley, S. H. A. 559
Howarth, E. 213
Jahoda, G. 203
Kempen, G. 149

PAGE
Kumar, D. 391
Lally, M. 17
Launay, G. 191
Lee, M. 137
Lind, P. M. 413
Lohr, J. M. 353
McAleavey, C. 137
McAskie, M. 243
McCusker, L. 507
Mohan, V. 391
Morris, P. E. 161
Morton, J. 31
Myers, D. H. 407
Nettlebeck, T. 17
Newstead, S. E. 165
Porter, L. 521
Proctor, J. 63
Reckase, M. D. 569
Renner, V. J. 1
Richardson, J. T. E. 315, 487
Ridgway, J. 161
Rimé, B. 507
Rowell, J. A. 1
Salter, D. 339
Somekh, D. E. 529
Springer, G. 339
Surtees, P. G. 367
Sweller, J. 553
Sykes, D. H. 429
Sykes, R. N. 515
Thornton, E. W. 359
Underwood, G. 327
Warden, D. A. 101
Warrington, E. K. 73
Wason, P. C. 479
Wetherick, N. E. 231, 501
Wolff, J. G. 377
Young, B. 89
Zangwill, O. L. 301

INDEX OF PUBLICATIONS REVIEWED

Argyle, M. & Beit-Hallahmi, B. *The Social Psychology of Religion* 286
Barnett, S. A. *The Rat. A Study in Behavior* 604
Belson, W. A. *Juvenile Theft: The Causal Factors* 446
Bench, R. J. *et al. Sound Perception in Mammals* 606
Botwinick, J. & Storandt, M. *Memory, Related Functions and Age* 121
Brown, B. B. *New Mind, New Body* 455
Brown, J. (ed.). *Recall and Recognition* 607
Burt, C. *The Gifted Child* 288
Canter, D. (ed.). *Environmental Interaction* 603
Cattell, R. B. & Child, D. *Motivation and Dynamic Structure* 602
Cheshire, N. M. *The Nature of Psychodynamic Interpretation* 601
CIBA Foundation Symposium 33. *Parent–Infant Interaction* 445
Clare, A. *Psychiatry in Dissent. Controversial Issues in Thought and Practice* 595
Clarke, A. M. & Clarke, A. D. B. (eds). *Mental Deficiency: The Changing Outlook* 443
Cohen, J. & Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioural Sciences* 452
Cooper, C. L. (ed.). *Theories of Group Processes* 451
Davidoff, J. B. *Differences in Visual Perception* 130
Dimond, S. J. & Beaumont, J. G. (eds). *Hemisphere Function in the Human Brain* 113
Dohrenwend, B. S. & Dohrenwend, B. P. (eds). *Stressful Life Events. Their Nature and Effects* 278
Doob, L. W. *Pathways to People* 596
Eacker, J. N. *Problems of Philosophy and Psychology* 455
Ehrenzweig, A. *The Psychoanalysis of Artistic Vision and Hearing. An Introduction to a Theory of Unconscious Perception* 281
Ekman, P. & Friesen, W. V. *Unmasking the Face* 454
Eliot, J. & Salkind, N. J. (eds). *Children's Spatial Development* 448
Ellis, N. R. (ed.). *Aberrant Development in Infancy* 444
Evans, R. M. *The Perception of Color* 291
Francis, H. *Language in Childhood: Form and Function in Language Learning* 447
Frank, G. *Psychiatric Diagnosis: A Review of Research* 119
Friedman, M. I. *Rational Behaviour, An Explanation of Behavior that is Especially Human* 131
Garfield, S. L. *Clinical Psychology. The Study of Personality and Behaviour* 284
Glucksberg, S. & Danks, J. H. *Experimental Psycholinguistics: An Introduction* 449
Goergen, D. *The Sexual Celibate* 604
Hess, E. H. *The Tell-tale Eye* 290
Hirsch, S. R. & Leff, J. P. *Abnormalities in Parents of Schizophrenics* 118
Howells, J. G. (ed.). *Modern Perspectives in the Psychiatry of Old Age* 120
Johnson, F. N. (ed.). *Lithium Research and Therapy* 295
Kennedy, A. & Wilkes, A. (eds). *Studies in Long Term Memory* 122
Kiernan, C. C. & Woodford, F. P. (eds). *Behaviour Modification with the Severely Retarded. Study Group 8 of the Institute for Research into Mental and Multiple Handicap* 593
Kinsbourne, M. & Smith, W. L. *Hemispheric Disconnection and Cerebral Function* 282
Klatzky, R. L. *Human Memory: Structures and Processes* 450
Kuller, R. (ed.). *Architectural Psychology* 292
Lader, M. *The Psychophysiology of Mental Illness* 117
Laska, E. M. & Bank, R. (eds). *Safeguarding Psychiatric Privacy: Computer Systems and their Uses* 597
Levi, L. (ed.). *Society, Stress and Disease*, vol. 2. *Childhood and Adolescence* 276
Levi, L. & Andersson, L. *Psychosocial Stress: Population, Environment and Quality of Life* 453
Lewis, M. R. & Rosenblum, L. A. (eds). *The Origins of Fear* 279
Linford Rees, W. (ed.). *Anxiety Factors in Comprehensive Patient Care: Where General Medicine and Psychiatry Meet* 116
Maccoby, E. E. & Jacklin, C. N. *The Psychology of Sex Differences* 441
MacKintosh, N. J. *The Psychology of Animal Learning* 439
McClelland, D. C. *Power* 602
McLeish, J. *Soviet Psychology: History, Theory, Content* 293
Nutting, J. M., Snr. *The Illusion of Attitude Change: Towards a Response Contagion Theory of Persuasion* 275
Padgham, C. A. & Saunders, J. E. *The Perception of Light and Colour* 129
Pask, G. *The Cybernetics of Human Learning and Performance* 123
Pitt, B. *Psychogeriatrics: An Introduction to the Psychiatry of Old Age* 120
Rabbitt, P. M. A. & Dornic, S. (eds). *Attention and Performance V* 124
Rachman, S. J. & Philips, C. *Psychology and Medicine* 115
Rock, I. *An Introduction to Perception* 128
Rosenzweig, N. *Community Mental Health Programmes in England: An American View* 287
Rummel, R. J. *Understanding Conflict and War*, vol. 1. *The Dynamic Psychological Field* 285
Ryle, A. *Frames and Cages: Repertory Grid Explorations of Personal Space* 284
Selye, H. *Stress without Distress* 600

Simpson, L. L. (ed.). *Drug Treatment of Mental Disorders* 597

Singleton, W. T. & Spurgeon, P. (eds). *Measurement of Human Resources* 294

Stevens, S. S. *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects* 127

Tizard, J. *et al.* (eds). *Varieties of Residential Experience* 288

Venables, P. H. & Christie, M. J. *Research in Psychophysiology* 440

Wallach, H. *Hans Wallach on Perception* 605

Weimer, W. B. & Palermo, D. S. (eds). *Cognition and the Symbolic Processes* 289

Weiner, I. B. *Principles of Psychotherapy* 280

Widroe, H. J. (ed.). *Human Behaviour and Brain Function* 115

Wilson, E. O. *Sociobiology: The New Synthesis* 598

Wyer, R. S., Jr. *Cognitive Organisation and Change* 126

Yates, A. J. *Theory and Practice in Behaviour Therapy* 593

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
The Journal of Occupational Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *The Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1976 subscription prices, payable in advance are: £20.00 net per volume post free (US $60.00 in the USA and Canada) for *The British Journal of Psychology*; £16.00 net (US $48.00) for *The British Journal of Medical Psychology*; £13.00 net (US $39.00) for *The Journal of Occupational Psychology*; and £19.00 net (US $57.00) for *The British Journal of Social and Clinical Psychology*. Subscriptions to any of these journals may be sent to any bookseller or subscription agent, or to the Cambridge University Press, Bentley House, 200 Euston Road, London, NW1 2DB. American and Canadian inquiries should be addressed to the Cambridge University Press, 32 East 57th Street, New York, N.Y. 10022.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £15.00 per volume post free ($34.00 in the USA). Subscriptions should be sent to the Society's Agents, John Wright & Sons, Ltd., 42–44 Triangle West, Bristol BS8 1EX, or to any bookseller.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The subscription is £6.00 a year post free, payable in advance, or £2.00 each number. The subscription price in the USA is $20.00. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh EH7 5JX.

MANUSCRIPTS

for *The British Journal of Psychology* should be sent to

Professor A. D. B. Clarke, Department of Psychology, The University, Hull HU6 7RX;

for the *British Journal of Medical Psychology* should be sent to

Professor Arthur H. Crisp, Psychiatric Research Unit, Atkinson Morley's Hospital, Copse Hill, London SW20 0NE;

for *The Journal of Occupational Psychology* should be sent to

Professor B. Shackel, Department of Human Sciences, University of Technology, Loughborough, Leicestershire;

for *The British Journal of Mathematical and Statistical Psychology* should be sent to

Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster LA1 4YF;

for *The British Journal of Social and Clinical Psychology* should be sent to

Dr Halla Beloff (Social Psychology Editor), Department of Psychology, University of Edinburgh, 60 Pleasance, Edinburgh EH8 9TJ: or to
Dr H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester M20 8LR;

for *The British Journal of Educational Psychology* should be sent to

Professor N. J. Entwistle, Department of Educational Research, University of Lancaster, Lancaster.

Contributors are advised to consult *Suggestions to Authors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable price 50p net (USA $2.00) post free, from Cambridge University Press.

All manuscripts should be submitted in duplicate.

SEPARATES

Contributors receive on application fifty copies of their papers free. Additional copies of separates may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £18.00 (£16.00 if paid by Banker's Order). They are entitled to purchase any or all of the Society's Journals at special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, 18–19 Albemarle Street, London W1X 4DN.

The British Journal of Psychology *Volume 67, Part 4, November 1976*

CONTENTS

HINTON, JOHN AND CRASKE, BRIAN
A relationship between Eysenck's P scale and change in muscle action potentials with attention to perceptual tasks *page* 461

FRIGON, JEAN-YVES
Extraversion, neuroticism and strength of the nervous system 467

CRAWSHAW, MARTIN AND CRASKE, BRIAN
Oculomotor adaptation to prisms: complete transfer between eyes 475

EVANS, J. ST B. T. AND WASON, P. C.
Rationalization in a reasoning task 479

RICHARDSON, JOHN T. E.
Procedures for investigating imagery and the distinction between primary and secondary memory 487

DURNDELL, A. J. AND WETHERICK, N. E.
The relation of reported imagery to cognitive performance 501

RIMÉ, BERNARD AND MCCUSKER, LEX
Visual behaviour in social interaction: the validity of eye-contact assessments under different conditions of observation 507

SYKES, R. N.
A comment on Meddis' 'A simple two-group test for matched scores with unequal cell frequencies' 515

EVANS, J. ST B. T.
A critical note on Quinton and Fellows' observation of reasoning strategies 517

FELLOWS, BRIAN J.
The role of introspection in problem-solving research: a reply to Evans 519

PORTER, LAWRENCE
Rehearsal span versus memory span: what does this comparison tell us about verbal memory? 521

SOMEKH, DAVID E.
The effect of embedded words in a brief visual display 529

DAY, R. H. AND DICKINSON, R. G.
The components of the Poggendorff illusion 537

SWELLER, J.
The effect of task complexity and sequence on rule learning and problem solving 553

HENLEY, S. H. A.
Responses to homophones as a function of cue words on the unattended channel 559

AGRUSO, VICTOR M., JR AND RECKASE, MARK D.
The relative difficulty position function 569

BRÉE, DAVID S. AND COPPENS, GEORGE
The difficulty of an implication task 579

ANNETT, MARIAN
A coordination of hand preference and skill replicated 587

BOOK REVIEWS 593

OTHER PUBLICATIONS RECEIVED 609

EDITORIAL ACKNOWLEDGEMENT 615

© The British Psychological Society 1976